FORM PTO-1390 (REV. 10-2000)	U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY'S DOCKET NUMBER 20267P
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		U.S. APPLICATION NO. (If known, see 37 CFR 1.5) 09/806088
INTERNATIONAL APPLICATION NO. PCT/US99/23253	INTERNATIONAL FILING DATE 10/05/1999	PRIORITY DATE CLAIMED 10/09/1998

TITLE OF INVENTION
DELTA 6 FATTY DESATURASE

APPLICANT(S) FOR DO/EO/US
KONSTANTIN PETRUKHIN AND C. THOMAS CASKEY

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. [X] This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. [ ] This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. [ ] This is an express request to begin national examination procedures [35 U.S.C. 371(f)] at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. [X] A proper Demand for International Preliminary Examination was made and the US was elected by the expiration of the 19th month from the earliest claimed priority date (PCT Article 31).
5. [X] A copy of the International Application as filed [35 U.S.C. 371(c)(2)].
   a. [ ] is attached hereto (required only if not communicated by the International Bureau).
   b. [ ] has been communicated by the International Bureau.
   c. [X] is not required, as the application was filed in the United States Receiving Office (RO/US).
6. [ ] An English language translation of the International Application as filed [35 U.S.C. 371(c)(2)].
7. [ ] Amendments to the claims of the International Application under PCT Article 19 [35 U.S.C. 371(c)(3)].
   a. [ ] are attached hereto (required only if not communicated by the International Bureau).
   b. [ ] have been communicated by the International Bureau.
   c. [ ] have not been made; however, the time limit for making such amendments has NOT expired.
   d. [ ] have not been made and will not be made.
8. [ ] An English language translation of the amendments to the claims under PCT Article 19 [35 U.S.C. 371(c)(3)].
9. [X] An oath or declaration of the inventor(s) [35 U.S.C. 371(c)(4)].
10. [ ] An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 [35 U.S.C. 371(c)(5)].

Items 11 to 16 below concern other document(s) or information included:

11. [ ] An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. [ ] An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. [ ] A **FIRST** preliminary amendment.
    [ ] A **SECOND** or **SUBSEQUENT** preliminary amendment.
14. [ ] A substitute specification.
15. [ ] A change of power of attorney and/or address letter.
16. [ ] Other items or information:

EXPRESS MAIL CERTIFICATE
DATE OF DEPOSIT March 26, 2001
EXPRESS MAIL NO. EE392097377US
I HEREBY CERTIFY THAT THIS CORRESPONDENCE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE AS EXPRESS MAIL "POST OFFICE TO ADDRESSEE" BEFORE 5 P.M. ON THE ABOVE DATE IN AN ENVELOPE ADDRESSED TO ASSISTANT COMMISSIONER FOR PATENTS, WASHINGTON, D.C., 20231.
MAILED BY [signature]
DATE 3/26/01

U.S. APPLICATION NO. (If known, see 37 CFR 1.5)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
09/806088	PCT/US99/23253	20267P

17. ☒ The following fees are submitted:	CALCULATIONS	PTO USE ONLY
BASIC NATIONAL FEE [37 CFR 1.492(a)(1)-(5)]:		
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee [37 CFR 1.445(a)(2)] paid to USPTO and International Search Report not prepared by the EPO or JPO **$1,000.00**		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO **$860.00**		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee [37 CFR 1.445(a)(2)] paid to USPTO **$710.00**		
International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4) **$690.00**		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4) **$100.00**		
ENTER APPROPRIATE BASIC FEE AMOUNT =	$100.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date [37 CFR 1.492(e)].	$0.00	

Claims	Number Filed	Number Extra	Rate		
Total Claims	15 - 20 =	0	X **$18.00**	$0.00	
Independent Claims	3 - 3 =	0	X **$80.00**	$0.00	
Multiple dependent claim(s) (if applicable)		0	+ **$270.00**	$0.00	
TOTAL OF ABOVE CALCULATIONS =				$100.00	
☐ Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.					
SUBTOTAL =				$100.00	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date [37 CFR 1.492(f)]. +				$0.00	
TOTAL NATIONAL FEE =				$100.00	
Fee for recording the enclosed assignment [37 CFR 1.21(h)]. The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property. +					
TOTAL FEES ENCLOSED =				$100.00	
				Amount to be refunded	
				charged	

a. ☐ A check in the amount of $ _____ to cover the above fees is enclosed.

b. ☒ Please charge my Deposit Account No. 13-2755 in the amount of $100.00 to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to the Deposit Account No. 13-2755. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive [37 CFR 1.137(a) or (b)] must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

MERCK & CO., INC.
Patent Department, RY60-30
P.O. Box 2000
126 East Lincoln Avenue
Rahway, New Jersey 07065-0970

DATE: MARCH 26, 2001

PHONE #: (732) 594-6734

Joseph A. Coppola
SIGNATURE

Joseph A. Coppola
NAME

38,413
REGISTRATION NUMBER

TITLE OF THE INVENTION

DELTA 6 FATTY ACID DESATURASE

CROSS-REFERENCE TO RELATED APPLICATIONS

Not applicable.

STATEMENT REGARDING FEDERALLY-SPONSORED R&D

Not applicable.

REFERENCE TO MICROFICHE APPENDIX

Not applicable.

FIELD OF THE INVENTION

The present invention is directed to novel human DNA sequences encoding a delta 6 fatty acid desaturase, an enzyme involved in the synthesis of essential fatty acids.

BACKGROUND OF THE INVENTION

Essential fatty acids (EFAs) are polyunsaturated fatty acids that cannot be manufactured by mammals, yet are required for a number of important biochemical processes, and thus must be supplied in the diet. The most important dietary EFAs are linoleic acid and alpha-linolenic acid (ALA). These two EFAs undergo a number of biosynthetic reactions that convert them into various other EFAs. Figure 1 depicts the biosynthetic reactions involving the two groups of EFAs, the n-6 EFAs (linoleic acid derivatives) and the n-3 EFAs (ALA derivatives). EFAs are formed from linoleic acid and ALA by a series of alternating reactions involving the removal of two hydrogens coupled with the insertion of an additional double bond (desaturation) and the lengthening of the fatty acid chain by the addition of two carbons (chain elongation). The enzymes catalyzing the desaturations and elongations are thought to be the same for both groups of EFAs.

Among the more important unsaturated fatty acids are the delta 6 unsaturated fatty acids, which are involved in the maintenance of membrane structure and function, the regulation of cholesterol synthesis and transport, and the prevention

of water loss from the skin. Delta 6 unsaturated fatty acids also serve as precursors of the eicosanoids, including the prostaglandins and leukotrienes (Horrobin, 1992, Prog. Lipid Res. 31:163-194). The double bond at the 6 position of delta 6 unsaturated fatty acids is introduced by a class of enzymes known as delta 6 desaturases.

Deficiencies in linoleic acid and ALA derivatives have been associated with skin diseases, diabetic complications, inflammatory and autoimmune disorders, cardiovascular disorders, complications of viral infection, and retinal dysfunction. For example, a deficiency in gamma-linolenic acid (GLA), which is produced from linoleic acid by the action of the enzyme delta 6 desaturase, can arise from the decreased activity of this enzyme that occurs in aging, stress, diabetes, eczema, and some infections, or from increased catabolism of GLA due to oxidation or rapid cell division, as occurs in inflammation or cancer. Clinical trials have demonstrated that dietary GLA supplementation can be effective in treating a number of conditions that are associated with GLA deficiency, *e.g.*, atopic eczema, mastalgia, diabetic neuropathy, viral infections, and some forms of cancer (Horrobin, 1990, Rev. Contemp. Pharmacother. 1:1-45).

Delta 6 desaturase is an example of a fatty acid desaturase. Fatty acid desaturases are enzymes that introduce a double bond into the carbon chain of fatty acids. They play vital roles in the biosynthesis of polyunsaturated fatty acids, including the essential fatty acids. Fatty acid desaturases are present in soluble and membrane-associated forms and require electron donors (for example, cytochrome b5) for their functioning.

Delta 6 desaturases catalyze the rate-limiting steps in the biosyntheses of the linoleic and ALA group EFAs shown in Figure 1. End products of the linoleic acid pathway include the eicosanoids (prostaglandins and leukotrienes). The end product of the ALA pathway is docosahexaenoic acid (DHA), an important component of membranes in the vertebrate retina. DHA is highly specific for retina and represents more than 50% of the fatty acids in the rod outer segment (ROS). It appears that DHA is important in maintaining the normal structure and function of the retina (Anderson et al., 1992, Neurobiology of Essential Fatty Acids, Bazan et al., eds., Plenum Press, New York, pages 285-294). Increased dietary consumption of DHA and its precursor, eicosapentaenoic acid, from seal meat and fish has been

linked to an increased incidence of macular degeneration in Greenland Eskimos (Rosenberg, 1987, Arct. Med. Res. 46:64-70).

Certain delta 6 desaturases have been cloned from plants. For example, a delta 6 desaturase has been cloned from borage (Sayanova et al., 1997, Proc. Natl. Acad. Sci. USA 94:4211-4216). This delta 6 desaturase is unusual in that its cytochrome b5 electron donor is present as an N-terminal extension of the enzyme rather than being synthesized as a separate protein. The borage delta 6 desaturase has been shown to be functional, in that transfer of the cloned gene encoding it to tobacco results in the synthesis of high levels of GLA and octadecatetraenoic acid (OTA) in the transgenic tobacco leaves. GLA and OTA are the products of delta 6 desaturase activity on linoleic acid and ALA, respectively.

Based on its hydropathy profile, the borage delta 6 desaturase appears to be a membrane-bound protein. Examination of the amino acid sequence of the borage enzyme, as well as the amino acid sequences of membrane-bound desaturases from a wide variety of organisms, has revealed three regions of conserved short motifs containing histidine residues ($HX_{(3\text{ or }4)}H$, $HX_{(2\text{ or }3)}HH$, and $HX_{(2\text{ or }3)}HH$) having a conserved spacing from each other (Shanklin et al., Biochemistry, 1994, 33:12787-12794).

A DNA sequence has been isolated from sunflower embryos that, judging from its sequence, appears to encode a delta 6 desaturase having a cytochrome b5-like moiety fused to its N-terminus (Sperling et al., 1995, Eur. J. Biochem. 232:798-805).

SUMMARY OF THE INVENTION

The present invention is directed to novel human DNA sequences that encode a delta 6 fatty acid desaturase, cytochrome b5-related protein (CYB5RP). The present invention includes genomic CYB5RP DNA as well as cDNA that encodes the CYB5RP protein. The genomic CYB5RP DNA is substantially free from other nucleic acids and has the nucleotide sequence shown in SEQ.ID.NO.:1. The cDNA encoding CYB5RP protein is substantially free from other nucleic acids and has the nucleotide sequence shown in SEQ.ID.NO.:2. Also provided is CYB5RP protein encoded by the novel DNA sequences. The CYB5RP protein is substantially free from other proteins and has the amino acid sequence shown in SEQ.ID.NO.:3.

Methods of expressing CYB5RP protein in recombinant systems are provided. Also provided are methods of producing delta 6 unsaturated fatty acids using DNA encoding CYB5RP or using CYB5RP protein.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts the enzymatic conversions involved in the linoleic acid (n-3) and alpha-linolenic acid (n-6) pathways of essential fatty acid synthesis.

Figure 2A-G shows the genomic DNA sequence of the CYB5RP gene (SEQ.ID.NO.:1). Underlined nucleotides in capitals represent exons. The start ATG codon at position 544 in exon 1 and the stop TGA codon at position 18,103 in exon 12 are shown in bold. The putative polyadenylation signal ATTAAA located approximately 20 base pairs upstream of the polyA tail is shown in bold italics (position 18,373 in exon 12). DNA sequence upstream of exon 1 represents a putative promoter region of the CYB5RP gene., as indicated by the presence of the TATA box at position 353 (underlined bold)..

Figure 3A-C shows the cDNA sequence (SEQ.ID.NO.:2) and the amino acid sequence (SEQ.ID.NO.:3) of CYB5RP. The region encompassing amino acids 1-102 represents the cytochrome b5 domain. The region encompassing amino acids 182-186 represents HIS BOX 1. The region encompassing amino acids 219-223 represents HIS BOX 2. The region encompassing amino acids 383-387 represents HIS BOX 3.

Figure 4 shows a portion of the cDNA sequence (SEQ.ID.NO.:4) and a portion of the amino acid sequence (SEQ.ID.NO.:5) of mouse CYB5RP.

Figure 5A shows a Kyte-Doolittle hydropathy plot of CYB5RP. Figure 5B shows the proposed membrane topology of CYB5RP based on its hydropathy plot. This membrane topology is similar to that proposed for other membrane-bound fatty acid desaturases (Shanklin et al., Biochemistry, 1994, 33:12787-12794). The amino acids shown in Figure 5B are portions of (SEQ.ID.NO.:3).

Figure 6 shows the output of the Profilescan program from the Wisconsin GCG package. The upper amino acid sequence is from CYB5RP (positions 31-78 of SEQ. ID. NO.3). The lower amino acid sequence is positions 1-48 of the cytochrome b5 profile (SEQ. ID. NO.:6.). The output shows that CYB5RP

contains a profile typical for the heme-binding domain of the cytochrome b5 protein family. Importantly, the region of identity includes the invariant HPGG motif, where histidine represents a heme axial ligand for iron.

Figure 7A and B show the results of BlastP searches of the GenBank database using the full-length CYB5RP amino acid sequence as the query. Figure 7A shows the hit with highest homology, a hypothetical protein from sunflower. The sunflower protein and CYB5RP share three His boxes (boxed) in which the spacing between the His boxes is conserved. Also boxed is the HPGG motif typical for the heme-binding domain of the cytochrome b5 protein family. In both proteins the first histidine of the third His box is replaced by glutamine (a typical feature of desaturases with delta 6 specificity). The upper amino acid sequences shown are from CYB5RP and are portions of SEQ. ID. NO.3. The lower amino acid sequences shown are portions of the amino acid sequence of the hypothetical protein from sunflower (Sperling et al., 1995, Eur. J. Biochem. 232:798-805). The sequence shown as positions 348-432 is SEQ. ID. NO.:7. The sequence shown as positions 22-74 is SEQ. ID. NO.:8. The sequence shown as positions 152-227 is SEQ. ID. NO.:9. Figure 7B shows the hit with the second highest homology, a delta 6 desaturase from *Borago oficinalis* (Sayanova et al., 1997, Proc. Natl. Acad. Sci. USA 94:4211-4216). The *Borago* protein and CYB5RP also share three His boxes with conserved spacing, as well as the HPGG motif. In both proteins the first histidine of the third His box is replaced by glutamine (a typical feature of desaturases with delta 6 specificity). The upper amino acid sequences shown are from CYB5RP and are portions of SEQ. ID. NO.3. The lower amino acid sequences shown are portions of the amino acid sequence of the *Borago* delta 6 desaturase. The sequence shown as positions 338-424 is SEQ. ID. NO.:10. The sequence shown as positions 12-64 is SEQ. ID. NO.:11. The sequence shown as positions 153-220 is SEQ. ID. NO.:12.

Figure 8 shows additional results of BlastP searches of the GenBank database using the CYB5RP protein as the query. Figure 8 shows the amino acid alignment between the CYB5RP protein and a delta 6 desaturase from *Synechocystis sp.* (strain pcc 6803) performed by the BlastP program. The *Synechocystis* delta 6 desaturase and CYB5RP share three His boxes, two of which are shown in Figure 8 (boxed). In both proteins the first histidine of the third His box is replaced by glutamine (a typical feature of desaturases with delta 6 specificity). The CYB5RP

sequence shown is a portion of SEQ. ID. NO.3. The *Synechocystis* sequence shown is SEQ. ID. NO:13.

Figure 9A shows the expression pattern of the CYB5RP gene in 9 human tissues, as determined by RT-PCR amplification with 21 cycles. Expression is detected in human retina, kidney, pancreas, placenta, and brain. Figure 9B shows the results of the analogous experiments performed with 25 cycles of amplification. Expression of the CYB5RP gene is seen in all the human tissues studied.

DETAILED DESCRIPTION OF THE INVENTION

For the purposes of this invention:

"Substantially free from other proteins" means at least 90%, preferably 95%, more preferably 99%, and even more preferably 99.9%, free of other proteins. Thus, a CYB5RP protein preparation that is substantially free from other proteins will contain, as a percent of its total protein, no more than 10%, preferably no more than 5%, more preferably no more than 1%, and even more preferably no more than 0.1%, of non-CYB5RP proteins. Whether a given CYB5RP protein preparation is substantially free from other proteins can be determined by such conventional techniques of assessing protein purity as, *e.g.*, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) combined with appropriate detection methods, *e.g.*, silver staining or immunoblotting.

"Substantially free from other nucleic acids" means at least 90%, preferably 95%, more preferably 99%, and even more preferably 99.9%, free of other nucleic acids. Thus, a CYB5RP DNA preparation that is substantially free from other nucleic acids will contain, as a percent of its total nucleic acid, no more than 10%, preferably no more than 5%, more preferably no more than 1%, and even more preferably no more than 0.1%, of non-CYB5RP nucleic acids. Whether a given CYB5RP DNA preparation is substantially free from other nucleic acids can be determined by such conventional techniques of assessing nucleic acid purity as, *e.g.*, agarose gel electrophoresis combined with appropriate staining methods, *e.g.*, ethidium bromide staining, or by sequencing.

"Substantially the same biological activity as CYB5RP" means being able to introduce a double bond into the 6 position of linoleic acid under conditions in which CYB5RP is able to introduce a double bond into the 6 position of linoleic acid.

A "conservative amino acid substitution" refers to the replacement of one amino acid residue by another, chemically similar, amino acid residue. Examples of such conservative substitutions are: substitution of one hydrophobic residue (isoleucine, leucine, valine, or methionine) for another; substitution of one polar residue for another polar residue of the same charge (*e.g.*, arginine for lysine; glutamic acid for aspartic acid); substitution of one aromatic amino acid (tryptophan, tyrosine, or phenylalanine) for another.

The present invention relates to the identification and cloning of cytochrome b5-related protein (CYB5RP), a gene which encodes a human delta 6 fatty acid desaturase. The gene is present on PAC clones 759J12, 756B3, 519O13, and 466A11 from an area of human chromosome 11q12 that has been shown to contain a gene related to Best's macular dystrophy (Cooper *et al.*, 1997, Genomics 41:185-192; Stöhr *et al.*, 1997, Genome Res. 8:48-56; Graff *et al.*, 1997, Hum. Genet. 101: 263-279). This linkage between the chromosomal location of the CYB5RP gene and the location of the gene related to Best's macular dystrophy can be used diagnostically by identifying restriction fragment length polymorphisms (RFLPs) in the vicinity of the CYB5RP gene, *e.g.*, in SEQ.ID.NO.:1. Such RFLPs will be associated with the Best's macular dystrophy gene and thus can be used to identify individuals carrying disease-causing forms of the Best's macular dystrophy gene.

CYB5RP was identified as an EST hit in sequence scanning data from PAC clones from human chromosome 11q12. In addition, a full length cDNA of CYB5RP was recovered from a human retina cDNA library. The genomic region of CYB5RP has been sequenced and the exon/intron organization of CYB5RP has been determined. The CYB5RP gene has 12 exons. The promoter region of CYB5RP was identified upstream of the 5' UTR by detecting consensus elements required for eukaryotic transcription. The expression pattern of CYB5RP was determined by RT-PCR analysis in 9 human tissues. The CYB5RP gene is expressed predominantly in human retina, kidney, pancreas, and placenta; lower levels of expression are also detected in brain,, heart, lung, liver, and skeletal muscle. Bioinformatic analysis revealed significant homology to a group of plant and bacterial fatty acid desaturases. All of the typical amino acid motifs present in these fatty acid desaturases are also present in CYB5RP. Kyte-Doolittle algorithm analysis predicts a transmembrane organization typical of fatty acid desaturases for CYB5RP (see Figure 5). CYB5RP is

unusual in that it contains a cytochrome b5 region in its N terminus. While many fatty acid desaturases utilize cytochrome b5 as an electron donor, most have not incorporated this cytochrome as part of their polypeptide chain.

That CYB5RP is a fatty acid desaturase is shown by the following evidence:

(1) CYB5RP possesses significant homology to a group of plant and microbial fatty acid desaturases;

(2) Like other fatty acid desaturases, CYB5RP has three conserved histidine boxes, with correct spacing between the boxes; and

(3) The predicted membrane topology of CYB5RP is similar to that of known fatty acid desaturases.

That CYB5RP is a delta 6 fatty acid desaturase is shown by the following evidence:

(1) CYB5RP contains a cytochrome b5-like moiety fused to its N-terminus. The only two fatty acid desaturases that contain cytochrome b5-like moiety fused to their N-termini are known or suspected to be delta 6 desaturases.

(2) The only two plant desaturases that are known or suspected to introduce a double bond in the 6 position have an atypical His box 3 (QI/LEHH), with a Q in the first position rather than an H. CYB5RP has the same atypical His Box 3.

(3) The only bacterial desaturase that is known to introduce a double bond in the 6 position has an atypical His box 3 (QVTHH), with a Q in the first position rather than an H. CYB5RP has the same atypical His Box 3.

CYB5RP is a target for the development of drugs for the treatment of disorders of lipid metabolism and for the treatment of conditions that require the modulation of the biosynthesis of prostaglandins and leukotrienes (asthma, pain, *etc.*). CYB5RP is also a target for the development of drugs for use in treating skin diseases, diabetic complications, reproductive disorders, including breast pain and premenstrual syndrome, inflammatory and autoimmune disorders, cardiovascular disorders, complications of viral infections, and various forms of retinal degeneration, including age-related macular degeneration.

CYB5RP is homologous to a delta 6 desaturase from *Borago oficinalis* (see Figure 7B). Both CYB5RP and this *Borago* delta 6 desaturase, unlike desaturases from higher plants, are unusual in containing a cytochrome b5-like

domain fused to their N-termini (Sayanova et al., 1997, Proc. Natl. Acad. Sci. USA 94:4211-4216; hereinafter "Sayanova"). The *Borago* desaturase has been expressed in transgenic tobacco, resulting in high levels of delta 6 desaturated fatty acids in the transgenic tobacco leaves, including high levels of γ-linolenic acid (GLA) (Sayanova). Given the medical importance of GLA, Sayanova proposed that transgenic plants, expressing the *Borago* delta 6 desaturase, would be valuable as sources of GLA. Similarly, CYB5RP, expressed in transgenic plants, is expected to provide a valuable source of GLA.

The present invention provides DNA encoding CYB5RP that is substantially free from other nucleic acids. The present invention also provides recombinant DNA molecules encoding CYB5RP. The present invention provides DNA molecules substantially free from other nucleic acids comprising the nucleotide sequence shown in Figure 2 as SEQ.ID.NO.:1. Analysis of SEQ.ID.NO.:1 revealed that this genomic sequence defines a gene having 12 exons. These exons collectively have an open reading frame that encodes a protein of 445 amino acids. When an alternatively spliced exon 8 is used, a CYB5RP protein of 433 amino acids, lacking amino acids 317-328, is produced. Thus, the present invention includes two cDNA molecules, encoding two forms of CYB5RP protein, that are substantially free from other nucleic acids. The first cDNA is shown in Figure 3 and has the nucleotide sequence SEQ.ID.NO.:2. The second cDNA is identical to the first, except that it does not contain the nucleotides at positions 1,019-1,054.

The present invention includes DNA molecules substantially free from other nucleic acids comprising the coding region of SEQ.ID.NO.:2. Accordingly, the present invention includes DNA molecules substantially free from other nucleic acids having a sequence comprising positions 71-1,405 of SEQ.ID.NO.:2. The present invention also includes DNA molecules substantially free from other nucleic acids having a sequence comprising positions 71-1,405 of SEQ.ID.NO.:2, except that the nucleotides at positions 1,019-1,054 are missing. Also included in the present invention are recombinant DNA molecules having a nucleotide sequence comprising positions 71-1,405 of SEQ.ID.NO.:2 and recombinant DNA molecules having a nucleotide sequence comprising positions 71-1,405 of SEQ.ID.NO.:2 with the exception that positions 1,019-1,054 are missing.

The novel DNA sequences of the present invention encoding CYB5RP, in whole or in part, can be linked with other DNA sequences, *i.e.*, DNA sequences to which CYB5RP is not naturally linked, to form "recombinant DNA molecules" encoding CYB5RP. Such other sequences can include DNA sequences that control transcription or translation such as, *e.g.*, translation initiation sequences, promoters for RNA polymerase II, transcription or translation termination sequences, enhancer sequences, sequences that control replication in microorganisms, sequences that confer antibiotic resistance, or sequences that encode a polypeptide "tag" such as, *e.g.*, a polyhistidine tract or the myc epitope. The novel DNA sequences of the present invention can be inserted into vectors such as plasmids, cosmids, viral vectors, P1 artificial chromosomes, or yeast artificial chromosomes.

Included in the present invention are DNA sequences that hybridize to at least one of SEQ.ID.NOs.:1 or 2 under stringent conditions. By way of example, and not limitation, a procedure using conditions of high stringency is as follows: Prehybridization of filters containing DNA is carried out for 2 hr. to overnight at 65°C in buffer composed of 6X SSC, 5X Denhardt's solution, and 100 µg/ml denatured salmon sperm DNA. Filters are hybridized for 12 to 48 hrs at 65°C in prehybridization mixture containing 100 µg/ml denatured salmon sperm DNA and 5-20 X 10^6 cpm of ^{32}P-labeled probe. Washing of filters is done at 37°C for 1 hr in a solution containing 2X SSC, 0.1% SDS. This is followed by a wash in 0.1X SSC, 0.1% SDS at 50°C for 45 min. before autoradiography.

Other procedures using conditions of high stringency would include either a hybridization carried out in 5XSSC, 5X Denhardt's solution, 50% formamide at 42°C for 12 to 48 hours or a washing step carried out in 0.2X SSPE, 0.2% SDS at 65°C for 30 to 60 minutes.

Reagents mentioned in the foregoing procedures for carrying out high stringency hybridization are well known in the art. Details of the composition of these reagents can be found in, *e.g.*, Sambrook, Fritsch, and Maniatis, 1989, Molecular Cloning: A Laboratory Manual, second edition, Cold Spring Harbor Laboratory Press. In addition to the foregoing, other conditions of high stringency which may be used are well known in the art.

The degeneracy of the genetic code is such that, for all but two amino acids, more than a single codon encodes a particular amino acid. This allows for the

construction of synthetic DNA that encodes the CYB5RP protein where the nucleotide sequence of the synthetic DNA differs significantly from the nucleotide sequence of SEQ.ID.NO.:2, but still encodes the same CYB5RP protein shown as SEQ.ID.NO.:3. Such synthetic DNAs are intended to be within the scope of the present invention. Also with the scope of the present invention are synthetic DNAs that encode a CYB5RP protein lacking amino acids 317-328 of SEQ.ID.NO.:3.

Another aspect of the present invention includes host cells that have been engineered to contain and/or express DNA sequences encoding CYB5RP protein. Such recombinant host cells can be cultured under suitable conditions to produce CYB5RP protein. An expression vector containing DNA encoding CYB5RP protein can be used for expression of CYB5RP protein in a recombinant host cell. Recombinant host cells may be prokaryotic or eukaryotic, including but not limited to, bacteria such as *E. coli*, fungal cells such as yeast, mammalian cells including, but not limited to, cell lines of human, bovine, porcine, monkey and rodent origin, plant cells such as tobacco, and insect cells including but not limited to *Drosophila* and silkworm derived cell lines. Cell lines derived from mammalian species which are suitable for recombinant expression of CYB5RP protein and which are commercially available, include but are not limited to, L cells L-M(TK$^-$) (ATCC CCL 1.3), L cells L-M (ATCC CCL 1.2), 293 (ATCC CRL 1573), Raji (ATCC CCL 86), CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC CRL 1658), HeLa (ATCC CCL 2), C127I (ATCC CRL 1616), BS-C-1 (ATCC CCL 26) and MRC-5 (ATCC CCL 171).

A variety of mammalian expression vectors can be used to express recombinant CYB5RP in mammalian cells. Commercially available mammalian expression vectors which are suitable include, but are not limited to, pMC1neo (Stratagene), pSG5 (Stratagene), pcDNAI and pcDNAIamp, pcDNA3, pcDNA3.1, pCR3.1 (Invitrogen), EBO-pSV2-neo (ATCC 37593), pBPV-1(8-2) (ATCC 37110), pdBPV-MMTneo(342-12) (ATCC 37224), pRSVgpt (ATCC 37199), pRSVneo (ATCC 37198), and pSV2-dhfr (ATCC 37146). Following expression in recombinant cells, CYB5RP can be purified by conventional techniques to a level that is substantially free from other proteins. A description of vectors that can be used to express CYB5RP can be found in, *e.g.*, Goeddel, ed., 1990, Meth. Enzymol. vol. 185 or Perbal, 1988, A Practical Guide to Molecular Cloning, John Wiley and Sons, Inc.

The present invention includes CYB5RP protein substantially free from other proteins. The amino acid sequence of the full-length CYB5RP protein is shown in Figure 3 as SEQ.ID.NO.:3. Thus, the present invention includes CYB5RP protein substantially free from other proteins having the amino acid sequence SEQ.ID.NO.:3. Also included in the present invention is a CYB5RP protein that is produced from an alternatively spliced CYB5RP mRNA where the protein has the amino acid sequence of SEQ.ID.NO.:3 with the exception that amino acids 317-328 are missing.

As with many proteins, it is possible to modify many of the amino acids of CYB5RP and still retain substantially the same biological activity as the original protein. Thus, the present invention includes modified CYB5RP proteins which have amino acid deletions, additions, or substitutions but that still retain substantially the same biological activity as CYB5RP. It is generally accepted that single amino acid substitutions do not usually alter the biological activity of a protein (see, *e.g.*, Molecular Biology of the Gene, Watson *et al.*, 1987, Fourth Ed., The Benjamin/Cummings Publishing Co., Inc., page 226; and Cunningham & Wells, 1989, Science 244:1081-1085). Accordingly, the present invention includes polypeptides where one amino acid substitution has been made in SEQ.ID.NO.:3 wherein the polypeptides still retain substantially the same biological activity as CYB5RP. The present invention also includes polypeptides where two or more amino acid substitutions have been made in SEQ.ID.NO.:3 wherein the polypeptides still retain substantially the same biological activity as CYB5RP. In particular, the present invention includes embodiments where the above-described substitutions are conservative substitutions. In particular, the present invention includes embodiments where the above-described substitutions do not occur in the His boxes of CYB5RP. In particular, the present invention includes embodiments where the above-described substitutions do not occur in positions where the amino acid present in those positions in CYB5RP is the same as the amino acid present in the corresponding position of the sunflower protein depicted in Figure 1 of Sperling et al., 1995, Eur. J. Biochem. 232:798-805 when these two proteins are aligned by BLASTP analysis. In particular, the present invention includes embodiments where the above-described substitutions do not occur in positions where the amino acid present in those positions in CYB5RP is the same as the amino acid present in the corresponding position of the

CCCTCTACCCCTGTCCCATCAGGC (SEQ.ID.NO.:15)

One of skill in the art would recognize that many other primer pairs based upon SEQ.ID.NO.:2 would also be suitable.

PCR reactions can be carried out with a variety of thermostable enzymes including but not limited to AmpliTaq, AmpliTaq Gold, or Vent polymerase. For AmpliTaq, reactions can be carried out in 10 mM Tris-Cl, pH 8.3, 2.0 mM $MgCl_2$, 200 μM for each dNTP, 50 mM KCl, 0.2 μM for each primer, 10 ng of DNA template, 0.05 units/μl of AmpliTaq. The reactions are heated at 95°C for 3 minutes and then cycled 35 times using the cycling parameters of 95°C, 20 seconds, 62°C, 20 seconds, 72°C, 3 minutes. In addition to these conditions, a variety of suitable PCR protocols can be found in PCR Primer, A Laboratory Manual, edited by C.W. Dieffenbach and G.S. Dveksler, 1995, Cold Spring Harbor Laboratory Press; or PCR Protocols: A Guide to Methods and Applications, Michael *et al.*, eds., 1990, Academic Press .

A suitable cDNA library from which a clone encoding CYB5RP can be isolated would be Human Retina 5'-stretch cDNA library in lambda gt10 or lambda gt11 vectors (catalog numbers HL1143a and HL1132b, Clontech, Palo Alto, CA). The primary clones of such a library can be subdivided into pools with each pool containing approximately 20,000 clones and each pool can be amplified separately.

By this method, a cDNA fragment encoding an open reading frame of either 445 amino acids (SEQ.ID.NO.:3) or an open reading frame of 433 amino acids (SEQ.ID.NO.:3 lacking the amino acids at positions 317-328) can be obtained. This cDNA fragment can be cloned into a suitable cloning vector or expression vector. For example, the fragment can be cloned into the mammalian expression vector pcDNA3.1 (Invitrogen, San Diego, CA). CYB5RP protein can then be produced by transferring an expression vector encoding CYB5RP or portions thereof into a suitable host cell and growing the host cell under appropriate conditions. CYB5RP protein can then be isolated by methods well known in the art.

As an alternative to the above-described PCR method, a cDNA clone encoding CYB5RP can be isolated from a cDNA library using as a probe oligonucleotides specific for CYB5RP and methods well known in the art for screening cDNA libraries with oligonucleotide probes. Such methods are described

in, *e.g.*, Sambrook *et al.*, 1989, Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory, Cold Spring Harbor, New York; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K., Vol. I, II. Oligonucleotides that are specific for CYB5RP and that can be used to screen cDNA libraries can be readily designed based upon the cDNA sequence of CYB5RP shown in SEQ.ID.NO.:2 and can be synthesized by methods well-known in the art.

Genomic clones containing the CYB5RP gene can be obtained from commercially available human PAC or BAC libraries available from Research Genetics, Huntsville, AL. PAC clones containing the CYB5RP gene (*e.g.*, PAC clones 759J12, 756B3, 519O13, and 466A11) are commercially available from Research Genetics, Huntsville, AL (Catalog number for individual PAC clones is RPCI.C). Alternatively, one may prepare genomic libraries, especially in P1 artificial chromosome vectors, from which genomic clones containing the CYB5RP can be isolated, using probes based upon the CYB5RP sequences disclosed herein. Methods of preparing such libraries are known in the art (Ioannou *et al.*,1994, Nature Genet. 6:84-89).

The present invention also provides oligonucleotide probes, based upon SEQ.ID.NO.:2 that can be used to determine the level of CYB5RP RNA in a sample. In particular, the present invention includes DNA oligonucleotides comprising at least 18 contiguous nucleotides of SEQ.ID.NO.:2. Also provided by the present invention are corresponding RNA oligonucleotides. The DNA or RNA oligonucleotide probes can be packaged in kits.

In addition to the utilities described above, the present invention makes possible the recombinant expression of the CYB5RP protein in various cell types. In particular, it is advantageous to recombinantly express CYB5RP in plant cells. Such expression in plant cells provides a method for the production of high levels of valuable EFAs such as GLA and OTA in the recombinant plant cells. An example of such recombinant expression of a delta 6 fatty acid desaturase, in that case from borage, is described in Sayanova et al., 1997, Proc. Natl. Acad. Sci. USA 94:4211-4216 (Sayanova). The recombinant expression of the borage delta 6 desaturase led to the production of high levels of GLA and OTA in the leaves of the tobacco plants in which it was expressed. The procedures described in Sayanova can be easily adapted to express CYB5RP in tobacco, thus providing an additional, useful way to produce

large amounts of valuable EFAs. Known methods of recombinantly expressing genes in other plant species beside tobacco can be used to express CYB5RP in those other species.

The present invention also makes possible the development of assays which measure the biological activity of the CYB5RP protein. Such assays using recombinantly expressed CYB5RP protein are especially of interest.

Assays for CYB5RP protein activity can be used to screen libraries of compounds or other sources of compounds to identify compounds that are activators or inhibitors of the activity of CYB5RP protein. Such identified compounds can serve as "leads" for the development of pharmaceuticals that can be used to modulate the activity of CYB5RP in patients suffering from conditions where that activity is abnormal, *e.g.*, skin diseases, diabetic complications, inflammatory and autoimmune disorders, cardiovascular disorders, complications of viral infection, and retinal dysfunction such as macular degeneration.

Such assays may comprise:

(a) recombinantly expressing CYB5RP protein in a host cell;

(b) measuring the biological activity of the recombinantly expressed CYB5RP protein in the presence and in the absence of a substance suspected of being an activator or an inhibitor of CYB5RP protein;

where a change in the biological activity of the recombinantly expressed CYB5RP protein in the presence as compared to the absence of the substance indicates that the substance is an activator or an inhibitor of CYB5RP protein.

In particular embodiments, the biological activity of the recombinantly expressed CYB5RP protein is the ability to introduce a double bond into the 6 position of linoleic acid or alpha-linoleic acid.

In some embodiments, it may be advantageous to insert additional steps between steps (a) and (b). Such additional steps might include lysing the host cell and fractionating its contents in order to partially purify the recombinantly expressed CYB5RP, thus facilitating exposure of the recombinantly expressed CYB5RP to the substance as well as to any substrate used in the assay.

The present invention includes activators and inhibitors identified by the methods described herein as well as pharmaceutical compositions comprising

such activators and inhibitors. The activators and inhibitors are generally combined with pharmaceutically acceptable carriers before use to form pharmaceutical compositions. Examples of such carriers and methods of formulation of pharmaceutical compositions containing activators or inhibitors and carriers can be found in Remington's Pharmaceutical Sciences. To form a pharmaceutically acceptable composition suitable for effective administration, such compositions will contain an effective amount of the activator or inhibitor.

Therapeutic or prophylactic compositions are administered to an individual in amounts sufficient to treat or prevent conditions where CYB5RP activity is abnormal. The effective amount can vary according to a variety of factors such as the individual's condition, weight, sex and age. Other factors include the mode of administration. The appropriate amount can be determined by a skilled physician.

Compositions can be used alone at appropriate dosages. Alternatively, co-administration or sequential administration of other agents can be desirable.

The compositions can be administered in a wide variety of therapeutic dosage forms in conventional vehicles for administration. For example, the compositions can be administered in such oral dosage forms as tablets, capsules (each including timed release and sustained release formulations), pills, powders, granules, elixirs, tinctures, solutions, suspensions, syrups and emulsions, or by injection. Likewise, they can also be administered in intravenous (both bolus and infusion), intraperitoneal, subcutaneous, topical with or without occlusion, or intramuscular form, all using forms well known to those of ordinary skill in the pharmaceutical arts.

Advantageously, compositions can be administered in a single daily dose, or the total daily dosage can be administered in divided doses of two, three or four times daily. Furthermore, compositions can be administered in intranasal form via topical use of suitable intranasal vehicles, or via transdermal routes, using those forms of transdermal skin patches well known to those of ordinary skill in that art. To be administered in the form of a transdermal delivery system, the dosage administration will, of course, be continuous rather than intermittent throughout the dosage regimen.

The dosage regimen utilizing the compositions is selected in accordance with a variety of factors including type, species, age, weight, sex and medical condition of the patient; the severity of the condition to be treated; the route

of administration; the renal, hepatic and cardiovascular function of the patient; and the particular composition thereof employed. A physician or veterinarian of ordinary skill can readily determine and prescribe the effective amount of the composition required to prevent, counter or arrest the progress of the condition. Optimal precision in achieving concentrations of composition within the range that yields efficacy without toxicity requires a regimen based on the kinetics of the composition's availability to target sites. This involves a consideration of the distribution, equilibrium, and elimination of a composition.

The present invention also includes antibodies to the CYB5RP protein. Such antibodies may be polyclonal antibodies or monoclonal antibodies. The antibodies of the present invention are raised against the entire CYB5RP protein or against suitable antigenic fragments of the protein that are coupled to suitable carriers, *e.g.*, serum albumin or keyhole limpet hemocyanin, by methods well known in the art. Methods of identifying suitable antigenic fragments of a protein are known in the art. See, *e.g.*, Hopp & Woods, 1981, Proc. Natl. Acad. Sci. USA 78:3824-3828; and Jameson & Wolf, 1988, CABIOS (Computer Applications in the Biosciences) 4:181-186.

For the production of polyclonal antibodies, CYB5RP protein or an antigenic fragment, coupled to a suitable carrier, is injected on a periodic basis into an appropriate non-human host animal such as, *e.g.*, rabbits, sheep, goats, rats, mice. The animals are bled periodically and sera obtained are tested for the presence of antibodies to the injected antigen. The injections can be intramuscular, intraperitoneal, subcutaneous, and the like, and can be accompanied with adjuvant.

For the production of monoclonal antibodies, CYB5RP protein or an antigenic fragment, coupled to a suitable carrier, is injected into an appropriate non-human host animal as above for the production of polyclonal antibodies. In the case of monoclonal antibodies, the animal is generally a mouse. The animal's spleen cells are then immortalized, often by fusion with a myeloma cell, as described in Kohler & Milstein, 1975, Nature 256:495-497. For a fuller description of the production of monoclonal antibodies, see Antibodies: A Laboratory Manual, Harlow & Lane, eds., Cold Spring Harbor Laboratory Press, 1988.

Gene therapy may be used to introduce CYB5RP polypeptides into the cells of target organs, *e.g.*, the pigmented epithelium of the retina or other parts of the

retina. Nucleotides encoding CYB5RP polypeptides can be ligated into viral vectors which mediate transfer of the nucleotides by infection of recipient cells. Suitable viral vectors include retrovirus, adenovirus, adeno-associated virus, herpes virus, vaccinia virus, and polio virus based vectors. Alternatively, nucleotides encoding CYB5RP polypeptides can be transferred into cells for gene therapy by non-viral techniques including receptor-mediated targeted transfer using ligand-nucleotide conjugates, lipofection, membrane fusion, or direct microinjection. These procedures and variations thereof are suitable for *ex vivo* as well as *in vivo* gene therapy. Gene therapy with CYB5RP polypeptides will be particularly useful for the treatment of diseases where it is beneficial to elevate CYB5RP activity.

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are intended to fall within the scope of the appended claims.

Various publications are cited herein, the disclosures of which are incorporated by reference in their entireties.

WHAT IS CLAIMED:

1. A recombinant DNA molecule encoding a polypeptide having the amino acid sequence of SEQ.ID.NO.:3.

2. A recombinant DNA molecule comprising a nucleotide sequence selected from the group consisting of:
SEQ.ID.NO.:1;
SEQ.ID.NO.:2;
SEQ.ID.NO.:2 lacking positions 1,019-1,054;
positions 71-1,405 of SEQ.ID.NO.:2; and
positions 71-1,405 of SEQ.ID.NO.:2 lacking positions 1,019-1,054.

3. A DNA molecule that hybridizes under stringent conditions to the DNA molecule of claim 2.

4. An expression vector comprising the DNA of claim 1.

5. A recombinant host cell comprising the DNA of claim 1.

6. A CYB5RP protein, substantially free from other proteins, having an amino acid sequence selected from the group consisting of SEQ.ID.NO.:3 and SEQ.ID.NO.:3 lacking positions 317-328.

7. The CYB5RP protein of claim 6 containing a single amino acid substitution.

8. The CYB5RP protein of claim 7 where the substitution is a conservative substitution.

9. The CYB5RP protein of claim 6 containing amino acid substitutions where the substitutions do not occur in positions where the amino acid

present in CYB5RP at those positions is also present in the corresponding position in the delta 6 desaturase from sunflower when CYB5RP and the delta 6 desaturase from sunflower are aligned by BLASTP analysis or where the substitutions do not occur in positions where the amino acid present in CYB5RP at those positions is also present in the corresponding position in the delta 6 desaturase from *Synechocystis* when CYB5RP and the delta 6 desaturase from *Synechocystis* are aligned by BLASTP analysis or where the substitutions do not occur in positions where the amino acid present in CYB5RP at those positions is also present in the corresponding position in the delta 6 desaturase from borage when CYB5RP and the delta 6 desaturase from borage are aligned by BLASTP analysis.

10. An antibody that binds specifically to the CYB5RP protein of claim 6.

11. A DNA or RNA oligonucleotide probe comprising at least 18 contiguous nucleotides of at least one of the sequences of claim 2.

12. A method for determining whether a substance is an activator or an inhibitor of CYB5RP protein comprising:

(a) recombinantly expressing the CYB5RP protein of claim 6 in a host cell;

(b) measuring the biological activity of the recombinantly expressed CYB5RP protein in the presence and in the absence of a substance suspected of being an activator or an inhibitor of CYB5RP protein;

where a change in the biological activity of the recombinantly expressed CYB5RP protein in the presence as compared to the absence of the substance indicates that the substance is an activator or an inhibitor of CYB5RP protein.

13. The method of claim 12 where the biological activity of CYB5RP protein is the ability to introduce a double bond into the 6 position of linoleic acid.

14. A pharmaceutical composition comprising an activator or an inhibitor of CYB5RP.

15. A method of treating macular degeneration comprising administering to a patient an effective amount of the pharmaceutical composition of claim 14.

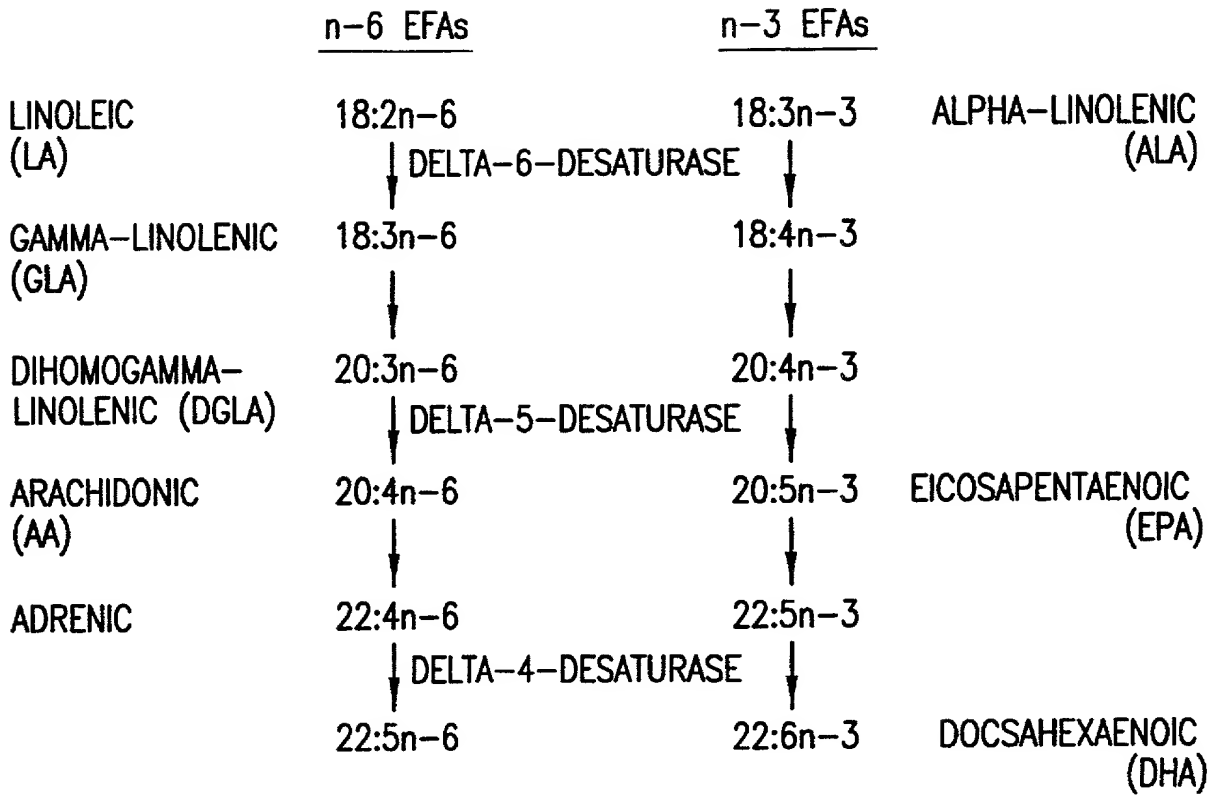

n-6 EFAs
n-3 EFAs
LINOLEIC (LA)
18:2n-6
18:3n-3
ALPHA-LINOLENIC (ALA)
DELTA-6-DESATURASE
GAMMA-LINOLENIC (GLA)
18:3n-6
18:4n-3
DIHOMOGAMMA-LINOLENIC (DGLA)
20:3n-6
20:4n-3
DELTA-5-DESATURASE
ARACHIDONIC (AA)
20:4n-6
20:5n-3
EICOSAPENTAENOIC (EPA)
ADRENIC
22:4n-6
22:5n-3
DELTA-4-DESATURASE
22:5n-6
22:6n-3
DOCSAHEXAENOIC (DHA)

FIG.1

```
   1  gctcacagac cgggactccg cctccggttc ccgagggcgt ggcgaggcgc
  51  tgcgggacgc ccaacaggtg cgtgttgtgt ccccaggccc cgcgctccgg
 101  gtggagtcaa gagcctggaa gccggcagcc cgggaaaagg gggcgggacg
 151  gtgccccggg gcagggctgg gtggcggccg ctgtcctccc gggaggggcg
 201  ggccgcctcg acgccgccct ccctggcggc caatggagac cgaggccccg
 251  cgcctggatt ggagcggacg cgggggtcag ccagccttgg gggccggggc
 301  ctggccgggg gcgggggggc aggcgaggcg aggcgggcgc cgtccgcgcg
 351  gttataaggc ggggagttcc ctgcgccgcg agccgggagg cgcacgctcg
 401  ctcgtacggc ggccgcggcg gcagggcggg gccggagcag cgggcggcgg
 451  cggaggcggc gcccgggagc gctCTTCGCT TCCCTCGGGG TCTTGCTCGG
 501  ACCTCGGCCA CCGCCTGGGA TCCCCAGGAC TCGTGCGTGC AGCATGGGCG
 551  GCGTCGGGGA GCCGGGACCG CGGGAGGGAC CCGCGCAGCC GGGGGCGCCG
 601  CTGCCCACCT TCTGCTGGGA GCAGATCCGC GCGCACGACC AGCCCGGCGA
 651  CAAGTGGCTG GTCATCGAGC GCCGCGTCTA CGACATCAGC CGCTGGGCAC
 701  AGCGGCACCC AGGGGGCAGC CGCCTCATCG GCCACCACGG CGCTGAGGAC
 751  GCCACGgtaa ggaagccata aggaagccac ccaccggcgg gtggagcctg
 801  gagctcggtc gtgggcgtga tgtcccgctc cacctgtggg gccttagcat
 851  cctccctccc ctcgctgacc tttgacctcc acgccgggac ccagagttgg
 901  ggtggactag ccagggccag atgtggggta gggagggcag ttccctgcgt
 951  ggaggacccg cagctgtcca cggagcaggt ctgcgggggа ggaggggccc
1001  tcagaggtgg gtgtgtcatg ctgcagagcc tgccctgggt gaggggctgc
1051  cctgttgctc ccaggtccct gtttcagttc tgggtcccca tgctgggtgc
1101  ttgctgagtg ctaggggtag ggcagggcag ggtccccagg ggccggtaag
1151  gacatgccat tagaggctgg gggctgggcc ggcctgaggt ctgtggcttt
1201  cccaagagct tctgtaaagg gctcagggac agtgactcac ctctccgggc
1251  tagcagctgc acgtgggagg gctttgccag ccaggctggg tgggcctctc
1301  ctggaagcac agtcacccca ggaacaggct ggcccctggg gaccccaact
1351  tcccaatccc agcccctgtc tagacaggca gggatgtagc ctggccccag
1401  ggtactgtct ggctggagtc cagtggtgga gcagcccgac cagccccttt
1451  tccttagtta cccacctgca taataggggt tggggccacg atgccctgtc
1501  cttgaccctc caaatttcta ggttggccac actgggtatc aggaaggtct
1551  tcaagacccg aggacatgaa tcctgaatgc tggctttttg ggcagcagcg
1601  gaggttctgt ccagtcccag gactgtcggc gtccctcttg ccagggccac
1651  ctgctctctg ccgattgcca tctccagcat gttggacaat cttcactgga
1701  ctctttgagg aagaaagccc ctcttttccc tttccacccc atgaagctga
1751  ggagtgagaa taagaatcct cctgaaattc taaaaaaaga aaaaaaaaaa
1801  aaagagaacg ccttgtccgt ggctgttcag gcgccagacg ctggcccgag
1851  gggacagcac agccgtggga tgaagcagcc tgggggcagt atttgagcgt
1901  gcaggtgttt gcatgtctgg gtgagtgtgg tgtgtgtgcc tgcctttctg
1951  ccagggcgtg gcgaggtgag gggcacggct tctccccaaa ggccttgctg
2001  agccctggcc tcccttcaag gagtcttgtg gatgcctgct ctggtctttt
2051  tttaaaaaag tatctatttt atttattatt atttgtttaa aaatagagac
2101  agggtctcac tatgttgctc gggctggtct caaagtcctg ggttcaagca
2151  ttcctcctgc ctcagcctcc gaaagttctg ggattacagg catgagccac
2201  cactcccggc ctgctctagt cttttgtaac ctagaggaca gtatggatac
2251  agaaaacttt actccccacc aaccgccgga gacagagtct tgctctgcca
2301  cccagactgg agtgcaatgg cgccatcttg gctcactgca acctccgcct
2351  cccaggttca agcgattctc ctgcctcagc ctcccgagta gctgggatta
2401  cgggcacgcg ccaccacgcc cagcatattg tatttttagt agagacgggg
2451  tttcaccatg ttggccaagc tggtctcgaa ctcctgacct cgtgatccac
2501  ccacctcggc ctcccaaagt gctgggatta caggcgtgag ccaccacgcc
2551  cggctgggat acagaaagct tttatttcat cactgtttcc tgcctggtgc
```

FIG.2A

```
2601  caggcccatg ctggggttcc tcccaagtgg aattactgac ttaacattta
2651  gcttgggatc ctgagacttc catcacacag ttttctcatt gattcgcagc
2701  caataatatc tgttttaaaa acatctcagg ccgagcgctg tggctcacac
2751  ctgtaatccc agcactttgg gaggctgagg tgggcagatc acctgaggtc
2801  gggagtttga gaccagcctg accaacatgg agaaaccctg tctcttctaa
2851  aaaaatacaa aattagccag gcgtggtggc gcatgcctgt aatcccagca
2901  ctttgggagg ctgaggcagg agaatcgctt gaacccagga gacggaggtt
2951  ccggtgagcc gagatcgcgc cattgcactc cagcctgggc aacaagagca
3001  aaactccgtc tcaaacaaac aaacaaaaaa catctctctg ctccttgggg
3051  ccgggtgcca gctctgctat tggaggcact gagcgacctt gaagcaggca
3101  tgtcactcct ctgtgcccca gtttactcat ctgtaaagtg ggagagctgg
3151  ggcagacagt gagctggctg agggcaggac tgtgtctcct caagcccatg
3201  gcccagggct gccaggtagt agtttgtatt cggtaaatgc tgctggcccc
3251  taagtgtgag cgtgccctgc aaactgcagc gtatggtggg acagccctgc
3301  acggctaccc ctttcctggg tgaccttatt tggttacggt cctatctgaa
3351  gtaggaaagg gacactttag gctgtctctt agctccctca aggccccaca
3401  gcctggacta gagttgccag aaatacttgg tccattcagg ccaaagggac
3451  tgtgaggttg ctgggatggt gcaatcagtc tttgtccatg atgaacccac
3501  agggtagacc aggggttggg ccagcccagt gccctgtgta gttgagccca
3551  ggccccaggc atcccatccc gggcggtggc ctcaggtgga ggtggggcag
3601  ccagttgcca gggatgtgtt ccagcggtca cctctcacca gccccggctg
3651  cccatcagct gttctcaagt ccaggcaatg aagccttcct gccaggaaat
3701  tcccagagtt tctgtgccat gaagtcagcc tgtggccatc ttgggacaca
3751  aggccgggtg ccctggggag agtactctgg gcccttggcc aggtttgtct
3801  gagagtcata ggcagcctga tactagtgga gccagccagg gagggatgag
3851  gcccagccgc tgctggccat aagtatataa gggccatgtg ctgagtgcct
3901  actatgtgcc aggttttgaa atcagtactt gatttattga aaccctctct
3951  tttaatcctc aaggtgcccc tatgaggcac gtaccattta ttgttattgc
4001  cacttgacag atgagaaaac agaggctcag agaggcaaag tggcttgaaa
4051  ttcagtgatt ggtctgggat ttgaatccac agccatgttc ttaagggcat
4101  gctatgctgc cacctatcct gtttatttcc ggcactcatt gattcttcaa
4151  tgtttgactc attaaatcca tcagtgagca tcttctctgt gtcatgcatg
4201  gttctcacct ctgaagatgt agctgtgagc aaaacttcta cagggaatga
4251  gttcacagca gagggatcag ctagagcaaa ggctcagagg tgggaccgtg
4301  cgtcctgtgt tccaggaata cagtatggct gcagcagaga gcagtggaga
4351  gagggcctgg cagtgaggtc tagaggcggc cgggctggct catgctggat
4401  gtttgtgtcc tcggaaggac tttggcttta ttttaaagag gatggggagc
4451  cccagagagc acagcaggga agcctgggga gtctgatgga catttaaaag
4501  gatccttaat ggagagagtg aaggcagagc cttccagaag ggtaagagaa
4551  gggaggatgg agacctgccc tcccccaagg gaggccactc agaagaggta
4601  gagtgtggcc agggcagaga gcaagagagg ctgtggacac aggcacactg
4651  gtccagtgag agccattaga cacattagat ttagcttcat gttgtcttta
4701  gagagggagc cagcctggcc tcgctctatg atcttggaca catcctttca
4751  cttctgggtc tcagtttccc cattagtgtg atgaggatga gaatgctttt
4801  gtcctgggca cactatgagg gtggtgctgg gcacctgggt gcctggttac
4851  catgggcaac aaagctctat tcatgggtgt ggtgaatgca ttgcccacag
4901  caactcaggg cggatgagga gtttcccagc agcccctggt gccctttcgg
4951  ctgaagccct aacaactgtg ggaaaatcca agttccagca gacccctga
5001  gccctctgcc ttaggaccct ccttctaggt ggttctctga gcctggcctg
5051  agctggagga gggagtggcc agtgctgcag cagaggctgc ttcatagtaa
5101  ttgcagccaa cagttattga ctaggcactg ttctgagggg tttagatgtg
5151  gtaactgatt gaattcgcct aacaacttta tgaggtaagt cctattgtta
5201  gcccattttg tagatgagga gactgagttt gaaactgggg ggtgtaatgg
5251  aaccttctca ggacccttga agggtagggc ctttgtactc gggccacgag
```

FIG.2B

5301 ggtggggttt gtgtctgggt gggagctggg gagggacagg actaggatta
5351 ggcagatctg aggccacagg agttggttgg ggggtggctc cagagccact
5401 ccactccctc ctaccacatt gactgccttg aaagtcccct aatggccact
5451 cccatgaagt gtgactgctc tgggctcccc gcaggcgttt tctgcaaggc
5501 caccgcccac ccaggcccct tccccagagg ggctgcagtg ccttgctcct
5551 tccttgtggg aagagttggg attgtctggc gtcagcagga tactgccccc
5601 gggcatccct cccggtctct tcctgcgggt ttctgatgaa acagccaggc
5651 tccagtagtg gagccagagg tcagtggtgg agagaggacc aggagccaga
5701 gggtatagct gctttggggc tactgtgggg tcagggacac ttgtgaggcc
5751 aagcgtcctg gctgcaggag ccctcacata tatgcccacc cttcaccagg
5801 acattgaggg gtgctggggg acaggggtag ctttttgggg gtgtctgcct
5851 tcgacttggg ctccgctaca caggccaaat ttggatgtcc catgtttaga
5901 gctgtgtttc tttgggacct cttggggcct cagtttcctc atctgtaaaa
5951 tgggatactg atagtgcttc cccactggcc tcctctgacg ggcgccaggg
6001 agaggatggg acggagcatg gtgtgctggg cacgctcctg ctgtacccac
6051 ccacctggga gaggggagag gcaggaatgt cctgggggtg tcctttgagg
6101 catagccctg tcaccccaac atcctacaaa ggcatgagaa ggcagcgagg
6151 acagaccccg accacctgag ccctcagcag ccctgccaca ctccctgctt
6201 cacccccttc ctgactgatc tggcacattc ttgattctcc tagggagtga
6251 cccaaaatcc ctccctgccc tgctgtgtct ctggggtgga aggaggctgc
6301 cagcccctcc tctctcccag cctcaggctt ggccaggact taacaggcag
6351 gcagagaagc agcttctcca ctctcttccc tgacacctgt aggcccctcc
6401 tgcaggcact tacctctaag tggactctca ggaggaggct catcagggct
6451 gcagggctca gaaagagctg ggctgtggag ctcttgccaa ccgccaggcc
6501 ccttctaagt gctttagcgc caccgactgc atcctcccag cagccttgtg
6551 agatggggat ttgtggttcc cagtttactg atgagaaata ctgatgagag
6601 atgggtgtgg tcttgtctgg ggctccctgg ctcctggata gcagctcagg
6651 ttccatcctg ggcaggctgg ctctgggaca ccccccgac cagctgctgt
6701 gtgggattca cggtggggct tgggcagggc gtgggatctt ggggccaact
6751 gagccactct aggcttccag ggaccaaggc caggctgagc tgtctctgta
6801 tcctgagaga gcatgaacat cacagaagat gggcccgggt tcgaatccca
6851 gctctgccac tactaactgg gacctgggca ggggtccctt cccgctgagc
6901 cttcatttcc tcaccagcaa aatggttcgt gcccctgctt tgggggctgt
6951 ggagggttgg ctcttgtcta cttgttcata cctgctgttg agcagctgct
7001 ctgtgccggc ctctgaggat gccactgtga acagagcctg tcgctacctc
7051 caggagcttg tgtttagggg tgccgttttg attccagcac tttcacccag
7101 ctctgctccg gtacccgatg agagacgtcg agtgccgctt tccactcgct
7151 tgggtgcgtg tgggggttgg ggggacaggg ctttgtgcac gtagccctgg
7201 gtggatgttc ctgggtgcac ttagggtgtg tgagggtggg acctcccaca
7251 gttccctgag gctccactga tgaggtccaa gaaccgcctt cctgcccccc
7301 agcccaggct cccagcagct gggcccttgg cttcttgaga tagtgactgg
7351 cctcacggca aggacccccg cacaccacct aggagaactg ctgcttcccc
7401 tctgttccag gagtggcgac aagcacagtt tttcgctttt gtttttgttt
7451 tcttcacttt aagttccggg aaacgtgcag aatgtgcagg tttgttacat
7501 aggtatacat gtgccatggt ggtttgctgc acccgtcaac ccctcatcta
7551 ggttttaagc tccatataca ttaggcattt gtcctaatgc tctccctccc
7601 cttgcccctc acccgcccag taagccccgg tgtgtgatgt tcccttccct
7651 gtgtccatgt gttctcattg ttcaactctc acttatgagt gagaagagac
7701 ctggactctg atctaacctc ggtcaaatgg aactgtgtga ccttgaagaa
7751 gtagcttaac ctctctgagt cttagcttct gcctggcacc cccatcctta
7801 aggagaggcc cacagaggac caggtcacat gacctcagcc agttccagag
7851 aaggctgttt gcttccaggt ttcggcctga gtccaggccc ctgccctact
7901 cgcactccct gatagcatga gaagcacagc cccagggtgc ccacccagct
7951 ctgagagccc agcctgcttc ccagggaact gtcacagccc cacctgtccc

FIG.2C

```
 8001  ttccccagct ggagccctgt caatggcttt ggggttctct gacacagccc
 8051  tgaggggget cacacttccc cttatcattg caaggggtag atctggcttg
 8101  aaggccctgg ggcaggcttg gttctgtcct cccctgtcag tgcctcgaca
 8151  gggctggcct gggtgaatca ggaccaacgg gaaaggaggc gaggagacca
 8201  atctggaccc aagatcctca gctcaataag gtggccccag aactgacatg
 8251  gggtgataga gggaagggct gggagggagg agattctggg gccgcagcca
 8301  cagcttgcac gttgcgccgg gtgtgtctgt gcgtgccagc tgcatctttg
 8351  cgtaccatgt gtgcaaggct gtgtttggct gagtgttcat gtgggccgtg
 8401  attgtgggca tgtttctgag tgtctgagtg atgcctgctg gtgtgggctg
 8451  gtgggtgtgt ctgcatgtgc gtgtgtgtct ggggagtttc aaaggagaaa
 8501  gagggactca ccatcacgct ggctcagcct taaaaaggta ggacatcctg
 8551  acacgtgctg caacatggat ggaccttaag gacattgtgc tgagtgaaac
 8601  aagccagagg caaaggaaca aacatgtgat ttctcccaga tgaggtttcc
 8651  ggaggaggca gatctgtatg gacagaaggt agcatggtgg ttgccggggc
 8701  agggggagga gagaatggag aattagtgtt taatggggac agagtttcag
 8751  ttggggaagg tgaaaaggtt ctggagctgg atgatggtga tggttggaca
 8801  acactgtgca tgcacttaat accactgagc tggacaccta aaaatgctta
 8851  caatggtaaa tttcatgtat attttactac aatttttaaa aaattggctg
 8901  ggcgtggtgg cttatgcctg taatcccaac actttgggag gccaaggcgg
 8951  gaggattgct tgagctcagg agttcaacac cagcctgggc aatatggtga
 9001  aaccccgact ctacgaaata tacaaaaatt agcctggtgt ggtggcttgc
 9051  acctctaatc ccacctactc agtaggctaa ggcacaagaa tctcttgaac
 9101  ctgggaggtg gaggttgcag taagccgaga tcatgccact gcaacccagt
 9151  ctgggcgaca gagcaagact ctgtctcaaa aaataaaaga taaataaaaa
 9201  aattagaggc caggtgtggc tcacacctgt actctcaaca ctttgggagg
 9251  ctgaggtggg aggatcgctt gaagtcaggc atttaagaca tgcctaggca
 9301  acatagtgag accttgactc tacaaaaaaa ttcaaaagtt aatgagacat
 9351  ggtggcatgt gcctgtagtc ctagctgctg gggaggctga ggtgggagga
 9401  tcacttacga ccaggatttc aaggctgcag tgagctgtga ttgcatcact
 9451  gcactccagc ctggtgacag agtgaggccc tgtctcaaaa aaatttttca
 9501  gtgtttttct gggctgggcg tggtggctca ttcctgtaat tccagcactt
 9551  tgggaggctg aggtgggtgg attgcttgag cccaggagtt taagaccagc
 9601  tgggcaacat ggcaaacctc atctctacaa aaaataaaaa taaaaaatta
 9651  gctgggcatg gtggtgcaca cctgtactaa cagctacgag agaggctaag
 9701  gtgggaggat cacctgagcc cgggaggttg aggctgcagt gagccatgat
 9751  tgcaccactg cactctagcc tgggcgatac agcaagaccc tatctcaaaa
 9801  aaaaaaaaaa aaaaaaaaaa aaaaacaccc agtggggtca gtagaacccc
 9851  aagagtcttc ttccctccca gctcccctgt acaccagccc cagctctgca
 9901  ggtagctggg ggcccagaca gcttcctggg gacccccagc cttccctctg
 9951  ccctttttte taccagtttt gctgcccctc cttcaagact catgtccaga
10001  gggggtgaga tctgcactta tacagccccc tcctctgtaa tgagtgagcc
10051  aagtcagccc aggttattcc agaaggggca ccctaccagc cccccagtcc
10101  ccaagctgcc ctgggcctat aaaagcaggc aaggggaccc ctagtagatc
10151  atgtaggtgt tacctcttag tgggtgctgg aggggcctga agtgctttct
10201  tcccccaggg tggtaggaga atgtcctggc agtgacttca gggcccgctg
10251  tcacttccgt tttaagactc accagctggt aggctcatta gcaagaggac
10301  aataggaggc ccctgtcctc agtcagcttt cttcaaaggt gtttccttta
10351  gcaactggga ggcctccctt ctccagaccc atggggacaa caccacccag
10401  ctactggttc tataagctgc tgtatggctc tggctagccc attcagagaa
10451  agcctctgaa agtacaagga aaaaaatcag tccaagagct gtgaacaatt
10501  agtgagccga ttacaatacc aagaccacag gcagacctgg aaggctaagt
10551  gagcccaggt gtgaagttca agcttacttt acttctgggc cacttcctgg
10601  ctggtctctt tccctggccc ttatctttct cctggtctgt cttctcttct
10651  cacccccttt ctttactctt tcttccttct cctgcatcgt actccacccc
```

FIG.2D

10701 cactccagct attacacaga atcgcgagaa tgttggatta ttcattttat
10751 ttatgatgtt ttcttttttg taaaaataga gacaaggtct cactatgtgg
10801 cccaggctgg tcttgaactc ctggcctcaa gcaatcctcg tgccttggcc
10851 tcttacagtg ctgggattac agatgtgagc caccatgcct ggcccatttt
10901 atttacttta aaaaaaaaat taggctgggc gcggtggctc acacctataa
10951 ttccagcact ttgggaggcc aaggtgggca gatcaactga ggtcaggagt
11001 taaagaccag cttggccacc tggggtcagg agtttgagac cagctactcc
11051 ggaggctgag accggagaat tgcttgaacc caggaggtag aggttgcaat
11101 gaactgagat catgccattg catgccagcc tgggcaacag agcaagactg
11151 tctcaaaaaa aaaaaaaatt atgttttgtg ctcctgcttc ctgctttgta
11201 agtcaaatca gtttaactgt tcaagtgtct tccttgcaaa cccccaagga
11251 ctcaatgtgt gtcgcccttg actgatcccc ccgccccgtg acccagtggt
11301 cctcagttcc aggttttccc acctaccctt cacccactgc ttatgtttat
11351 aaaaacgggg taaatcaaat gttcgtgacc cagatcttat tctacatgca
11401 gtggaaactt gtatgactta agctttttgg aaaagcagaa cctttttcg
11451 tggttcaaga aatcaaagtc ttcccgggag gtctttctgt aaatccagag
11501 ctgcagatgt ttgaccgtgt tcagagaggg gcccttgtgc tgggtgaagt
11551 ggatggggca cagcaggcaa tgggtgaaaa gcaggacaac ctggggccct
11601 gggaggacca gggagggccc atgtctttga ctgttcatca gccggctgac
11651 ttcctgtccg cctgtcgtct gctctgccca tccatccgta gtccttccgc
11701 ctgtctctgc tggttgccgc tgtgctactc agctgtgtct gtctgtccgc
11751 ctgactgtct gctctccttc agGATGCCTT CCGTGCCTTC CATCAAGATC
11801 TCAATTTTGT GCGCAAGTTC CTACAGCCCC TGTTGATTGG AGAGCTGGCT
11851 CCGGAAGAAC CCAGCCAGGA TGGACCCCTG AATgtgagcc agagccctag
11901 gagaggctca gcccctgagg gagggggatg gctggagggc tgggagacat
11951 tgccacatgg ccaggagcag ctccctcggc attcgcccaa ggggatgcag
12001 agccagggct gagcctgccc tcccctccca gggggcaggc agttgaaagt
12051 gaagctgtag ggatgccctg agaagtccag ggctccagat ctggtttagc
12101 caggcactcg tttggatccc gaggcaagct ccctccctgt tgtcgcccag
12151 tgtccccatc aaaaggagga ttttgatgaa ctgatttctc tcctggctgt
12201 agcgtcttac ccaccccata ccttttggga gggagaggag gcttcaccac
12251 cagccagtgc tccagctcac accccgggct gggtactctt gtcacttcat
12301 tcctctttgc ccacacccct tgggcctggc gatgggagga gcggctgggg
12351 ctccaggaga atgggggtgg ggaggaattt cttccttggc tgatcggccc
12401 ctctgctatg gcagGCGCAG CTGGTCGAGG ACTTCCGAGC CCTGCACCAG
12451 GCAGCCGAGG ACATGAAGCT GTTTGATGCC AGTCCCACCT TCTTTGCTTT
12501 CCTACTGGGC CACATCCTGG CCATGGAGGT GCTGGCCTGG CTCCTTATCT
12551 ACCTCCTGGG TCCTGGCTGG GTGCCCAGTG CCCTGGCCGC CTTCATCCTG
12601 GCCATCTCTC AGgtgacccc agttctgtgt tgcagccacc ttaactgccc
12651 aacagacgtg ggcccccatg catctgggca ttgtgaacat atttgctaaa
12701 tgaatgaatg gacctatgaa aggatgaatg gatgaataaa cagatgaatg
12751 agtgaacagt ctgaaggccc atcaggcatg tctgtgggtc aagctgcatt
12801 ccagatgagc caagaagttc cttcttgaac agattccgat caagcacagg
12851 gccactgagc cagaggctgc tgccctgcag cttcatgaca cttacgagcc
12901 cctccacctc cctgggactc agttctcatc tgtaaaaaga ggacactggc
12951 ccacaagggt cttgaaatgg agcattagca cgggggtacc ctgcaagctg
13001 aaaggattca ctggggcccc aggccctggc gggctccgtc cttcccaaca
13051 gcttctgacc ctgcctctct ccccagGCTC AGTCCTGGTG TCTGCAGCAT
13101 GACCTGGGCC ATGCCTCCAT CTTCAAGAAG TCCTGGTGGA ACCACGTGGC
13151 CCAGAAGTTC GTGATGGGGC AGCTAAAGgt gagggtgggg tgggtggtca
13201 gccaggtgct gggtggcgct gggtctgccc aagtgtgtgg gcacagtcgg
13251 gggcacagcc tgccctgaga gccccctcct cctccacagG GCTTCTCCGC

FIG.2E

13301 CCACTGGTGG AACTTCCGCC ACTTCCAGCA CCACGCCAAG CCCAACATCT
13351 TCCACAAAGA CCCAGACGTG ACGGTGGCGC CCGTCTTCCT CCTGGGGGAG
13401 TCATCCGTCG AGgtgggtgg ggagggacct ggacaacctc tggctgggcc
13451 tgcagctgag ggggagctaa tgcactgggt ccccactctg cccctgacct
13501 agcccctgat ctggcctcca ctctggctgg gccaagctct gcccccgtgt
13551 ctttccttcc cacctcccaa cctgctgggg acgaccagcc cgcttgctag
13601 aatctagagt tgcctttgac ccttggcccc agccagcccc gtgaccttgc
13651 ccgggagaag gaggtggcct ggagagctgc tgtctccagc cgccgcctgt
13701 ctccacagTA TGGCAAGAAG AAACGCAGAT ACCTACCCTA CAACCAGCAG
13751 CACCTGTACT TCTTCCTGAg tgagtgtcca tctgtccttc tggggtgggg
13801 gagtgcctgg gcctgcactg tcctccctgc tgtcctggac cactcccagc
13851 cacttcctgg ggcggggcac gtctgtcagg tctccctggt catggcatcc
13901 tcccagcctc tgcagtctgt acacactctc ccagcagcat gcctttgccc
13951 cagctgtctc ccgtgcctgg gacaccttgc agccacgggc catcacagcc
14001 ctgctgggag cttccccaag ccccacgtag aatttcttct tgccctcact
14051 agagtggtcc ggagccctag agtctttggg cagttgttgg ggcggacaga
14101 gtgaggactc aagtctggcc ctgacttgcg gtgaagggtg gtgggaggtg
14151 gtggggtaag ggcagcctgg ggaggcttgg acacagaatt gggggtgata
14201 tggggtcatt cagctggatg tgaccagcac caacgtccca ggggcattcc
14251 tggagtaaca gagcccctca ctctggcgcc cactcacctt ggcagcccag
14301 ccccactcct gaacactctc atgccccttc ttgcagTCGG CCCGCCGCTG
14351 CTCACCCTGG TGAACTTTGA AGTGGAAAAT CTGGCGTACA TGCTGGTGTG
14401 CATGCAGTGG GCGgtgagtg gggttgccca ggaccccggg catacggctg
14451 ccgtggcagg aggtggtgcc tcgggggaca gtacctgccc atgaaggcaa
14501 acagggtgca catgtgcgtg caacagtgtg gctcacatgt atgcgtgcaa
14551 cagtgtggct cacatgtgtg cgcgcagcag gagagcgagt gtgcccgtga
14601 ctgtacgtgt ggtgggggggg ggttgaggaa caggggggt gtgggtctct
14651 ctcggtgagg gtgtcttccc aggaggagtt gctgggccga ctctgccagg
14701 catctgtgtc cctggcaggg tcttccccaa cacaccctgc atgacacctt
14751 cgtcactaaa atcagcctcg tgagctggca gggcaaggac cctgttcctt
14801 tactcagctg agaaaaccag agagggtggt ggcctgtcct gggctctgag
14851 gcaaatcagg cagaagggtt ggatgcctga ggtcctcctc ccacccacca
14901 ggcctccaga cctccgggca cctggagacc tctcggtatc gcctctgccc
14951 tcctctgcag GATTTGCTCT GGGCCGCCAG CTTCTATGCC CGCTTCTTCT
15001 TATCCTACCT CCCCTTCTAC GGCGTCCCTG GGGTGCTGCT CTTCTTTGTT
15051 GCTGTCAGgt atggcaggga gtggcgaggt cacacacagg cgacaggtga
15101 cccccactgc agccccccac cagagcttcc cttttcccgt ctgcagaatg
15151 gggccagtgg tactgcctcc ctggcttgct ggtggaatca cataaacaca
15201 agcgtggcag gagcccaggg tcggtgggtt tagggagcgt ggcctggctt
15251 gtaagtggcc cggtgggtgt cggagctgct ctggactcag cctcacagtg
15301 gacactgctc cattcagatt ctttaaacac tggcaagggg gcgatggcca
15351 caatcctatt gtacagataa ggaagtcaag gccacttggg gacagctgct
15401 ctccagcctc cactcagggt gcctaagtgg tgagctggac ctagggcagt
15451 gcccgagcct ccccacagGG TCCTGGAAAG CCACTGGTTC GTGTGGATCA
15501 CACAGATGAA CCACATCCCC AAGGAGATCG GCCACGAGAA GCACCGGGAC
15551 TGGGTCAGCT CTCAGgtggg cagcaggggt ggggcccatc ctgggtgggg
15601 tgggggtcc cagctaggag ccagatggca aagcagggat gaggccctga
15651 cggggctgcc aggtggggga tggtgccgtg gggtcaggga tctgcaacgg
15701 cctcctcaca tgtgccccgc cggcttccgg cagCTGGCAG CCACCTGCAA
15751 CGTGGAGCCC TCACTTTTCA CCAACTGGTT CAGCGGGCAC CTCAACTTCC
15801 AGATCGAGCA CCAgtgagtg tgggtgctgg gggccagtgg gaggtgggga
15851 gggggtcctg ggaggggatc ctgggagggg acccgtgggt ggggcctctc

FIG.2F

15901 tctggaatct cccacttcag gtgccagcat acgctcccca cccccagCCT
15951 CTTCCCCAGG ATGCCGAGAC ACAACTACAG CCGGGTGGCC CCGCTGGTCA
16001 AGTCGCTGTG TGCCAAGCAC GGCCTCAGCT ACGAAGTGAA GCCCTTCCTC
16051 ACCGCGCTGG TGGACATCGT CAGgtgaggc tgcagcccgg cccctctgtt
16101 ctggtggctt ccccagggcc tatgcctacc cttgtccagg tcagcctcat
16151 gctgagcccc cagggtccct gagcctttct gtccacgtcc catgcccttc
16201 ctcccttccc cagccttcac gcacacagtg agaatttctg gagcacctac
16251 tgcagactca caaacagcag tgcctgcggt gagcaggtct atgcaaacct
16301 acccccaaag gctgagggaa aaaagctaac agatccagtt tctcagaagg
16351 aaacacttaa cagggactca taaacagaag ccatgtctca gggccgggtg
16401 cggtggctca cgcctgtaat tccagcactt ggggaggctg aggtgggcgg
16451 atcacttgag gtcaggagtt cgagaccagc ctggccaaca tggtgaaacc
16501 ccgtctctac taaaaaaaaa aaaaaaaaac aaaacaaaac aaaaattagc
16551 tgggtgtggt ggcaggtgcc cataatccca gctacttggg aggctgaggg
16601 aggagaatca cttgaactcg caggggcaga ggttgcagtg agctgagatt
16651 gtgcctttgc agtccagcct gggcaacaga gcaagactct ctcaaaaaca
16701 aacaaaaaaa ccatgtctca ggcagccaag agttgggaca tcccctcaca
16751 cgccctctag aaagaaccct ctatatagca agcttttagg gtgaacccca
16801 tgcaggtggt tcttatgaac ctggtgacca ctggaggtta gataagcgtc
16851 tacaagagga ggttatctat gccatgagct tggcattcag ggtcaagcat
16901 cggtcatcag acagttttgc ttgaagatgg cattgccctt gtagcaatgc
16951 aggctctaga gagcttcctg ccctcttgga gctgatgttc cttccagcaa
17001 aggaaacagc aagcaattaa aataacaaat aagtacatta cagaagatgg
17051 gcaaaagaac aatgaaaagc ccctcagggt ggggacaggg gaggggaggg
17101 gggcggccag gcaggggcgg cagtttctaa ataggtggta gggtgggcag
17151 tattgacagg ctgacgtgtg agcagggaca gggaggaggg gagaggtctc
17201 gccacaggga catctggcaa agagcgttca ggcagagggc acttgaccct
17251 gaatgccaag ctcatggcat agatagccga ggcaggcatg caggcactca
17301 gagaagggac acgcccggct tgcatcttgg aaagctgccc ctactgggaa
17351 tgactggcgg gcaggagtcg aagtggaaaa ggagagcaga ggacactgca
17401 gccatccagg cgaggggtga tggggctcag cccttgtggt caccttggag
17451 gtggggaaca gaggccagat tccaggtctt atacctctgc gcctttgtac
17501 acgctgttcc ccttacttgg ttgcccttcc ttcctgtgct ggtgttcaga
17551 tgcccacttc tccttcatga tctctcccag cctgatgctc tgagcccctg
17601 ccatttggca cagcccttta gagcgcctgg cacagggctt cctagcagat
17651 tgttgacatt tctggctcca ctgcccaata tcaggcccaa gatcgggtgg
17701 gcaggttcca cgtcctctct gtccttgggt tgcagcgccc agcaggaggc
17751 agcaatggag aactgggtgc aggagggaca ggcccaccca ggctcatgcc
17801 tggacttggc cttggctgcc ctccagctcc cctacccgac acccgtcacc
17851 ccggtctaga ttccattcca gagaatgagc attcagctgt tctcccaacc
17901 caccctccag cccgcatcgc tgcctgcccc cagggaaggg aacccacagg
17951 gaatggggat ctccgctcac acttaccatg ggggatacag gggtgttagg
18001 atcttgcaac tgagctccta acacccaccc ccactgccac ccccacctcc
18051 cagGTCCCTG AAGAAGTCTG GTGACATCTG GCTGGACGCC TACCTCCATC
18101 AG**TGA**AGGCA ACACCCAGGC GGGCAGAGAA GGGCTCAGGG CACCAGCAAC
18151 CAAGCCAGCC CCCGGCGGGA TCGATACCCC CACCCCTCCA CTGGCCAGCC
18201 TGGGGGTGCC CTGCCTGCCC TCCTGGTACT GTTGTCTTCC CCTCGGCCCC
18251 CTCACATGTG TATTCAGCAG CCCTATGGCC TTGGCTCTGG GCCTGATGGG
18301 ACAGGGGTAG AGGGAAGGTG AGCATAGCAC ATTTTCCTAG AGCGAGAATT
18351 GGGGGAAAGC TGTTATTTTT AT***TATAAA***AT ACATTCAGAT GTATTATGGA
18401 GT

FIG.2G

```
  1  CTTCGCTTCCCTCGGGGTCTTGCTCGGACCTCGGCCACCGCCTGGGATCC   50

 51  CCAGGACTCGTGCGTGCAGCATGGGCGGCGTCGGGGAGCCGGGACCGCGG  100
  1                      M  G  G  V  G  E  P  G  P  R     10

101  GAGGGACCCGCGCAGCCGGGGGCACCGCTGCCCACCTTCTGCTGGGAGCA  150
 11  E  G  P  A  Q  P  G  A  P  L  P  T  F  C  W  E  Q   27

151  GATCCGCGCGCACGACCAGCCCGGCGACAAGTGGCTGGTCATCGAGCGCC  200
 28   I  R  A  H  D  Q  P  G  D  K  W  L  V  I  E  R  R  44

201  GCGTCTACGACATCAGCCGCTGGGCACAGCGGCACCCAGGGGGCAGCCGC  250
 45    V  Y  D  I  S  R  W  A  Q  R  H  P  G  G  S  R    60

251  CTCATCGGCCACCACGGCGCTGAGGACGCCACGGATGCCTTCCGTGCCTT  300
 61  L  I  G  H  H  G  A  E  D  A  T  D  A  F  R  A  F   77

301  CCATCAAGATCTCAATTTTGTGCGCAAGTTCCTACAGCCCCTGTTGATTG  350
 78   H  Q  D  L  N  F  V  R  K  F  L  Q  P  L  L  I  G  94

351  GAGAGCTGGCTCCGGAAGAACCCAGCCAGGATGGACCCCTGAATGCGCAG  400
 95    E  L  A  P  E  E  P  S  Q  D  G  P  L  N  A  Q   110

401  CTGGTCGAGGACTTCCGAGCCCTGCACCAGGCAGCCGAGGACATGAAGCT  450
111  L  V  E  D  F  R  A  L  H  Q  A  A  E  D  M  K  L  127

451  GTTTGATGCCAGTCCCACCTTCTTTGCTTTCCTACTGGGCCACATCCTGG  500
128   F  D  A  S  P  T  F  F  A  F  L  L  G  H  I  L  A  144

501  CCATGGAGGTGCTGGCCTGGCTCCTTATCTACCTCCTGGGTCCTGGCTGG  550
145   M  E  V  L  A  W  L  L  I  Y  L  L  G  P  G  W    160

551  GTGCCCAGTGCCCTGGCCGCCTTCATCCTGGCCATCTCTCAGGCTCAGTC  600
161  V  P  S  A  L  A  A  F  I  L  A  I  S  Q  A  Q  S  177

601  CTGGTGTCTGCAGCATGACCTGGGCCATGCCTCCATCTTCAAGAAGTCCT  650
178   W  C  L  Q  H  D  L  G  H  A  S  I  F  K  K  S  W  194

651  GGTGGAACCACGTGGCCCAGAAGTTCGTGATGGGGCAGCTAAAGGGCTTC  700
195   W  N  H  V  A  Q  K  F  V  M  G  Q  L  K  G  F    210
```

FIG.3A

```
701  TCCGCCCACTGGTGGAACTTCCGCCACTTCCAGCACCACGCCAAGCCCAA  750
211  S  A  H  W  W  N  F  R  H  F  Q  H  H  A  K  P  N   227

751  CATCTTCCACAAAGACCCAGACGTGACGGTGGCGCCCGTCTTCCTCCTGG  800
228   I  F  H  K  D  P  D  V  T  V  A  P  V  F  L  L  G  244

801  GGGAGTCATCCGTCGAGTATGGCAAGAAGAAACGCAGATACCTACCCTAC  850
245   E  S  S  V  E  Y  G  K  K  K  R  R  Y  L  P  Y   260

851  AACCAGCAGCACCTGTACTTCTTCCTGATCGGCCCGCCGCTGCTCACCCT  900
261  N  Q  Q  H  L  Y  F  F  L  I  G  P  P  L  L  T  L  277

901  GGTGAACTTTGAAGTGGAAAATCTGGCGTACATGCTGGTGTGCATGCAGT  950
278  V  N  F  E  V  E  N  L  A  Y  M  L  V  C  M  Q  W  294

951  GGGCGGATTTGCTCTGGGCCGCCAGCTTCTATGCCCGCTTCTTCTTATCC  1000
295   A  D  L  L  W  A  A  S  F  Y  A  R  F  F  L  S   310

1001 TACCTCCCCTTCTACGGCGTCCCTGGGGTGCTGCTCTTCTTTGTTGCTGT  1050
311  Y  L  P  F  Y  G  V  P  G  V  L  L  F  F  V  A  V  327

1051 CAGGGTCCTGGAAAGCCACTGGTTCGTGTGGATCACACAGATGAACCACA  1100
328   R  V  L  E  S  H  W  F  V  W  I  T  Q  M  N  H  I  344

1101 TCCCCAAGGAGATCGGCCACGAGAAGCACCGGGACTGGGTCAGCTCTCAG  1150
345   P  K  E  I  G  H  E  K  H  R  D  W  V  S  S  Q   360

1151 CTGGCAGCCACCTGCAACGTGGAGCCCTCACTTTTCACCAACTGGTTCAG  1200
361  L  A  A  T  C  N  V  E  P  S  L  F  T  N  W  F  S  377

1201 CGGGCACCTCAACTTCCAGATCGAGCACCACCTCTTCCCCAGGATGCCGA  1250
378   G  H  L  N  F  Q  I  E  H  H  L  F  P  R  M  P  R  394

1251 GACACAACTACAGCCGGGTGGCCCCGCTGGTCAAGTCGCTGTGTGCCAAG  1300
395   H  N  Y  S  R  V  A  P  L  V  K  S  L  C  A  K   410

1301 CACGGCCTCAGCTACGAAGTGAAGCCCTTCCTCACCGCGCTGGTGGACAT  1350
411  H  G  L  S  Y  E  V  K  P  F  L  T  A  L  V  D  I  427

1351 CGTCAGGTCCCTGAAGAAGTCTGGTGACATCTGGCTGGACGCCTACCTCC  1400
428   V  R  S  L  K  K  S  G  D  I  W  L  D  A  Y  L  H  444
```

FIG.3B

```
1401  ATCAGTGAAGGCAACACCCAGGCGGGCAGAGAAGGGCTCAGGGCACCAGC   1450
 445    Q                                                    445

1451  AACCAAGCCAGCCCCCGGCGGGATCGATACCCCCACCCCTCCACTGGCCA   1500

1501  GCCTGGGGGTGCACTGCCTGCCCTCCTGGTACTGTTGTCTTCCCCTCGGC   1550

1551  CCCCTCACATGTGTATTCAGCAGCCCTATGGCCTTGGCTCTGGGCCTGAT   1600

1601  GGGACAGGGGTAGAGGGAAGGTGAGCATAGCACATTTTCCTAGAGCGAGA   1650

1651  ATTGGGGGAAAGCTGTTATTTTTATATTAAAATACATTCAGATGTAAAAA   1700
```

FIG.3C

```
1   GTACAGCGGCAATGGGCGGTGTCGGGGAGCCCGGAGGGGGACTCGGGCCG   50
1                M  G  G  V  G  E  P  G  G  G  L  G  P    13

51  CGGGAGGGGCCCGCACCGCTGGGGGCGCCCCTACCCATCTTCCGCTGGGA   100
14  R  E  G  P  A  P  L  G  A  P  L  P  I  F  R  W  E    30

101 GCAGATCCGCCAGCATGACCTACCAGGCGACAAGTGGCTGGTCATCGAGC   150
31   Q  I  R  Q  H  D  L  P  G  D  K  W  L  V  I  E  R   47

151 GCCGTGTCTACGACATCAGCCGCTGGGCACAGCGGCACCCAGGGGGTAGC   200
48    R  V  Y  D  I  S  R  W  A  Q  R  H  P  G  G  S     63

201 CGCATCATCGGCCACCACGG  220
64  R  I  I  G  H  H     69
```

FIG.4

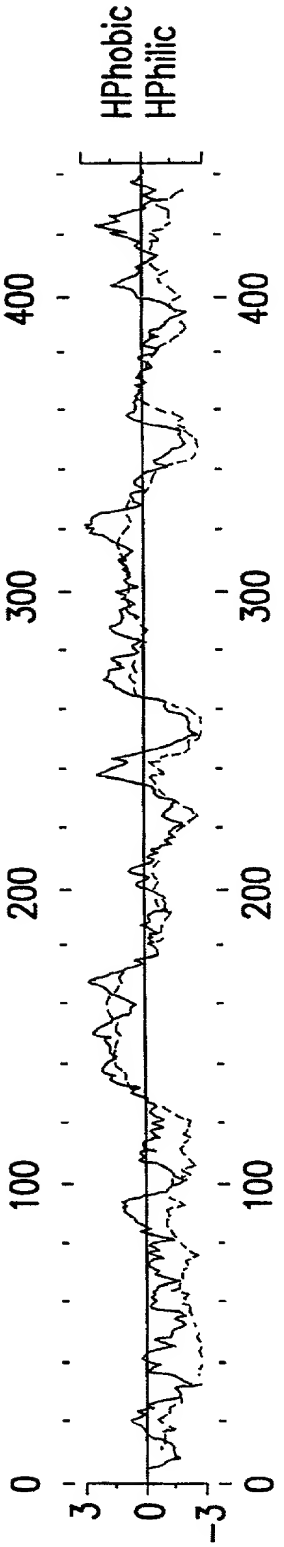

HPhobic
HPhilic
0
100
200
300
400
3
0
-3

FIG.5A

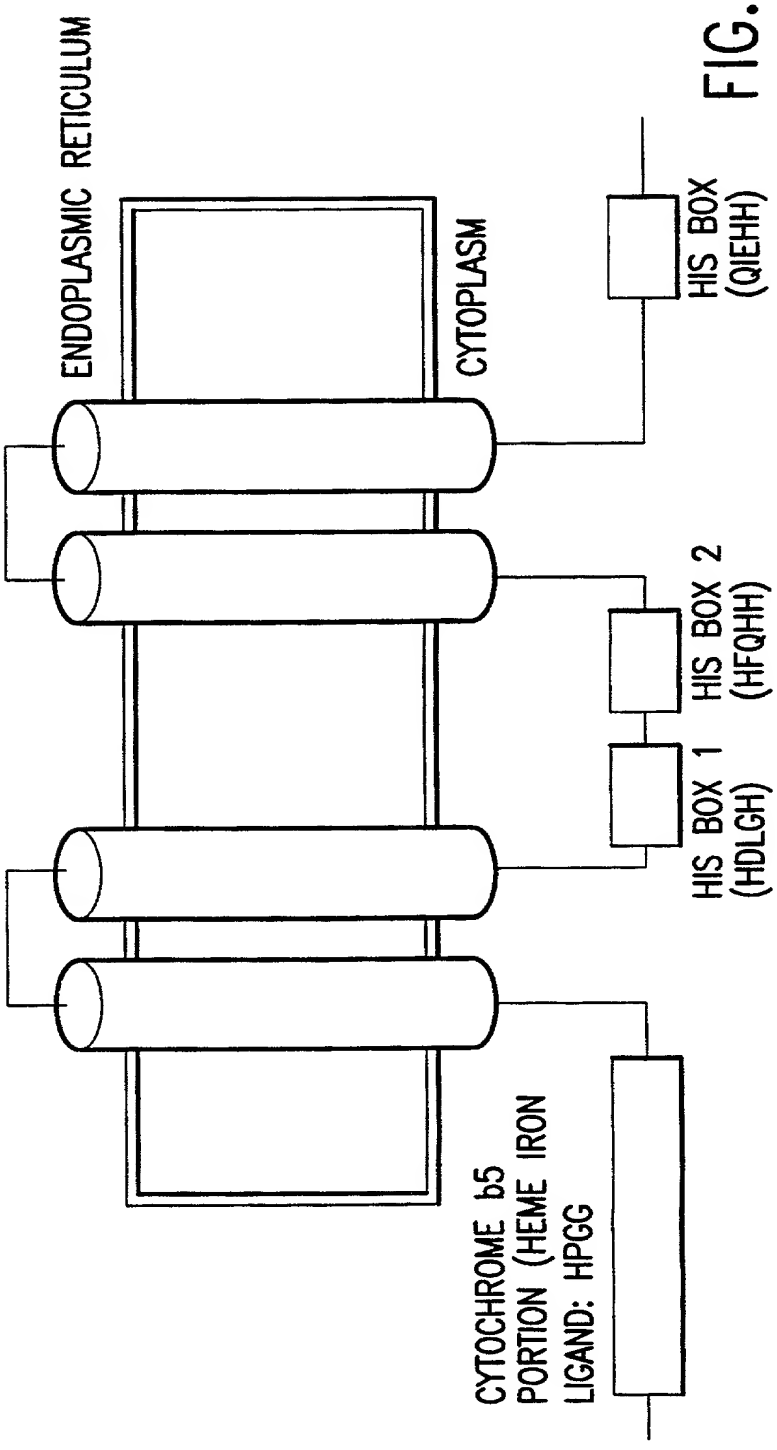

ENDOPLASMIC RETICULUM
CYTOPLASM
HIS BOX
(QIEHH)
HIS BOX 2
(HFQHH)
HIS BOX 1
(HDLGH)
CYTOCHROME b5
PORTION (HEME IRON
LIGAND: HPGG

FIG.5B

```
 PROFILESCAN of : CYB5rp_correct_protein check: 5714 from: 1 to: 445

 GETSEQ from bmd, December 2, 1997 14:20.

 Compare to profile library: GenRunData:profilescan.fil
..
-------------------------------------------------------------------------

 Profile: profiledir:cytochrome_b5.prf
   Gap weight:  4.50     Gap Length weight:  0.05
   Ave match:   0.27     Ave mismatch      : -0.21
(Peptide) PROFILEMAKE v4.40 of: 0191.Msf2{*} Length: 48
  Sequences: 24  MaxScore: 27.58  December 2, 1992 00:07
This profile is derived from PROSITE release 10.0 and has been tested
by a database search against SWISS-PROT release 26.0. A comparison
of the SWISS-PROT annotation and the results of the database search follows.
For further information about this motif, consult the . . .

Profile: profiledir:cytochrome_b5.prf alignment: 1

 Quality: 20.77    Gaps: 0
    Ratio: 0.43  Length: 48
 Normalized quality: 2.91

S     31 HDQPGDKWLVIERRVYDISRWAQRHPGGSRLIGHHGAEDATDAFRAFH 78
         |: ..: ||||. .|||:::| . ||||.  |  .||.|:||.| ::|
P      1 HNDGEETWLVVNGQVYDITKFLEEHPGGPDVIMEAAGTDATEEFEAIH 48

*************************************************************
*Cytochrome b5 family, heme-binding domain signature *
*************************************************************
```

FIG.6

pir:s68358 hypothetical protein - common sunflower
Length = 458

Score = 169 (79.4 bits), Expect = 2.8e-42, Sum P(4) = 2.8e-42
Identities = 31/85 (36%), Positives = 49/85 (57%)

His box 3

```
Query:  348 IGHEKHRDWVSSQLAATCNVEPSLFTNWFSGHLNFQIEHHLFPRMPRHNYSRVAPLVKSL 407
            +G  K  +W   Q   T ++  S + +WF G L FQ+EHHLFPR+PR +   ++P+ + L
Sbjct:  348 VGPPKGDNWFEKQTRGTIDIACSSWMDWFFGGLQFQLEHHLFPRLPRCHLRSISPICREL 407

Query:  408 CAKHGLSYEVKPFLTALVDIVRSLK 432
            C K+ L Y    F  A V  +++L+
Sbjct:  408 CKKYNLPYVSLSFYDANVTTLKTLR 432
```

Score = 133 (62.5 bits), Expect = 2.8e-42, Sum P(4) = 2.8e-42
Identities = 21/53 (39%), Positives = 35/53 (66%)

HPGG motif

```
Query:   26 EQIRAHDQPGDKWLVIERRVYDISRWAQRHPGGSRLIGHHGAEDATDAFRAFH 78
            ++++ H+ P D W+ I  +VY+++ WA+ HPGG   + +   +D TDAF AFH
Sbjct:   22 KELKKHNNPNDLWISILGKVYNVTEWAKEHPGGDAPLINLAGQDVTDAFIAFH 74
```

Score = 118 (55.5 bits), Expect = 2.8e-42, Sum P(4) = 2.8e-42
Identities = 25/76 (32%), Positives = 34/76 (44%)

His box 1 His box 2

```
Query:  165 LAAFILAISQAQSWCLQHDLGHASIFKKSWWNHVAQKFVMGQLKGFSAHWWNFRHFQHEA 224
            L+  IL ++  Q   L HD GH  +     WN  A  F+   + G S  WW + H  HH
Sbjct:  152 LSGAILGLAWMQIAYLGHDAGHYQMMATRGWNKFAGIFIGNCITGISIAWWKWTHNAHHI 211

Query:  225 KPNIFHKDPDVTVAPV 240
                N   DPD+   P+
Sbjct:  212 ACNSLDYDPDLQHLPM 227
```

Score = 34 (16.0 bits), Expect = 2.8e-42, Sum P(4) = 2.8e-42
Identities = 7/14 (50%), Positives = 9/14 (64%)

FIG. 7A

gp:bou79010_1 PID:g2062403 Borago officinalis delta 6 desaturase mRNA, complete cds. (gb:U79010) (NID:2062402)
Length = 448

Score = 179 (84.1 bits), Expect = 2.3e-42, Sum P(3) = 2.3e-42
Identities = 34/87 (39%), Positives = 48/87 (55%)

His box 3

```
Query:   348 IGHEKHRDWVSSQLAATCNVEPSLFTNWFSGHLNFQIEHHLFPRMPRHNYSRVAPLVKSL 407
             +G  K  +W   Q   T ++    + +WF G L FQIEHHLFP+MPR N  +++P V  L
Sbjct:   338 VGKPKGNNWFEKQTDGTLDISCPPWMDWFHGGLQFQIEHHLFPKMPRCNLRKISPYVIEL 397

Query:   408 CAKHGLSYEVKPFLTALVDIVRSLKKS 434
             C KH L Y    F  A     +R+L+ +
Sbjct:   398 CKKHNLPYNYASFSKANEMTLRTLRNT 424
```

Score = 144 (67.7 bits), Expect = 2.3e-42, Sum P(3) =2.3e-42
Identities = 23/53 (43%),Positives = 36/53 (67%)

HPGG MOTIF

```
Query:    26 EQIRAHDQPGDKWLVIERRVYDISRWAQRHPGGSRLIGHHGAEDATDAFRAFH 78
             ++++ HD+PGD W+ I+ + YD+S W + HPGGS  +     ++ TDAF AFH
Sbjct:    12 DELKNHDKPGDLWISIQGKAYDVSDWVKDHPGGSFPLKSLAGQEVTDAFVAFH 64
```

Score = 105 (49.3 bits), Expect = 2.3e-42, Sum P(3) = 2.3e-42
Identities = 22/68 (32%), Positives = 28/68 (41%)

His box 1 His box 2

```
Query:   176 QSWCLQHDLGHASIFKKSWWNHVAQKFVMGQLKGFSAHWWNFRHFQHHAKPNIFHKDPDV 235
             QS  + HD GH  +   S N     F    L G S  WW + H  HH   N    DPD+
Sbjct:   153 QSGWIGHDAGHYMVVSDSRLNKFMGIFAANCLSGISIGWWKWNHNAHHIACNSLEYDPDL 212

Query:   236 TVAPVFLL 243
                 p  ++
Sbject:  213 QVIPFLVV 220
```

FIG. 7B

pir:s35157 Delta(6)-desaturase - Synechocystis sp.
Length = 359

Score =126 (59.2 bits), Expect = 9.0e-09, Sum P(2) = 9.0e-09
Identities = 21/54 (38%), Positives = 33/54 (61%)

His box 3

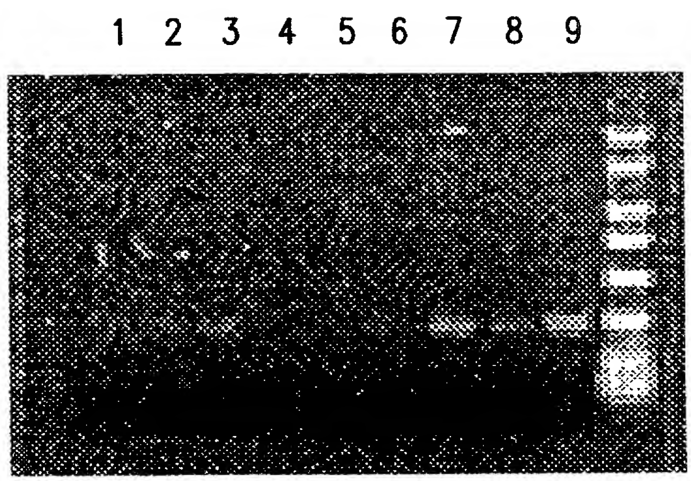
```
Query:   372 FTNWFSGHLNFQIEHHLFPRMPRHNYSRVAPLVKSLCAKHGLSYEVKPFLTALV 425
             F NMF G LN Q+ HHLFP +   +Y ++  ++K +C + G+ Y+V P    A +
Sbjct:   292 FWNWFCGGLNHQVTHHLFPNICHIHYPQLENIIKDVCQEFGVEYKVYPTFKAAI 345
```

Score = 36 (16.9 bits), Expect = 9.0e-09, Sum P(2) = 9.0e-09
Identities = 6/15 (40%), Positives = 8/15 (53%)

His box 2

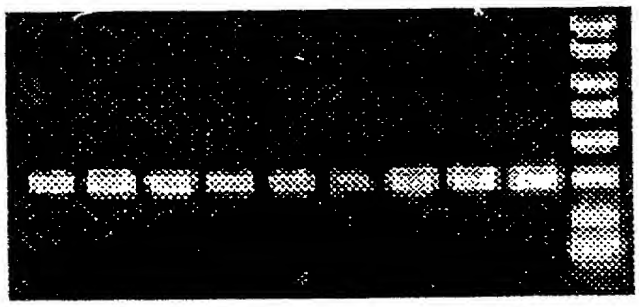
```
Query:   209 GFSAHWWNFRHFQHH 223
             G S+ W +RH   H
Sbjct:   113 GLSSFLWRYRHNYLH 127
```

FIG.8



1. Heart
2. Brain
3. Placenta
4. Lung
5. Liver
6. Skeletal Muscle
7. Kidney
8. Pancreas
9. Retina

FIG.9A



1 2 3 4 5 6 7 8 9 L PCR Marker

1. Heart
2. Brain
3. Placenta
4. Lung
5. Liver
6. Skeletal Muscle
7. Kidney
8. Pancreas
9. Retina

FIG.9B

Please type a plus sign (+) inside this box → [+]

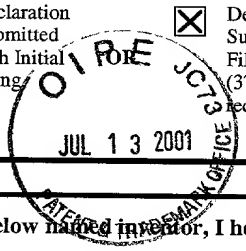
Approved for use through 9/30/2000. OMB 651-0032

SUBSTITUTE for PTO/SB/01 (12-97), DECLARATION FOR UTILITY OR DESIGN PATENT APPLICATION

DECLARATION AND POWER OF ATTORNEY FOR UTILITY OR DESIGN PATENT APPLICATION (37 CFR 1.63)

☐ Declaration Submitted with Initial Filing

☒ Declaration Submitted after Initial Filing (surcharge (37 CFR 1.16 (e)) required)



Attorney Docket Number	20267P
First Named Inventor	Petrukhin, et al.
COMPLETE IF KNOWN	
Application Number	09/806,088
Filing Date	March 26, 2001
Group Art Unit	
Examiner Name	

As a below named inventor, I hereby declare that:

My residence, post office address, and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

DELTA 6 FATTY DESATURASE

(Title of the Invention)

the specification of which

☐ is attached hereto

OR

☒ was filed on (MM/DD/YYYY) 03/26/2001 as United States Application Number or PCT International Application Number 09/806,088 and was amended on (MM/DD/YYYY) _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment specifically referred to above.

I acknowledge the duty to disclose to the Patent and Trademark Office all information known to me to be material to patentability as defined in 37 CFR 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. 119(a)-(d) or 365(b) of any foreign application(s) for patent or inventor's certificate, or 365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or of any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (MM/DD/YYYY)	Attorney Docket Number	Priority Claimed? YES	NO
PCT/US99/23253	PCT	10/09/1999	20267-PCT	☒	☐
				☐	☐
				☐	☐
				☐	☐

☐ Additional foreign application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto:

I hereby claim the benefit under 35 U.S.C. 119(e) of any United States provisional application(s) listed below.

Application Number(s)	Filing Date (MM/DD/YYYY)	Attorney Docket Number
60/103,760	10/09/1998	20267PV

Please type a plus sign (+) inside this box → [+]

Approved for use through 9/30/2000. OMB 651-0032

SUBSTITUTE for PTO/SB/01 (12-97), DECLARATION FOR UTILITY OR DESIGN PATENT APPLICATION

DECLARATION AND POWER OF ATTORNEY for Utility or Design Patent Application

I hereby claim the benefit under 35 U.S.C 120 of any United States application(s), or 365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of 35 U.S.C. 112, I acknowledge the duty to disclose information known to me to be material to patentability as defined in 37 CFR 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application or PCT Parent Application Number	Parent Filing Date (MM/DD/YYYY)	Parent Patent Number *(if applicable)*
60/103,760	10/09/1998	
pct/us99/23253	10/05/1999	

☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto.

As a named inventor, I hereby appoint, respectively and individually, as my attorneys or agents with full power of substitution and revocation, the following registered practitioner(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

☐ Customer Number [ ] → *Place Customer Number Bar Code Label here*

OR

☒ Registered practitioner(s) name/registration number listed below

Name	Registration Number	Name	Registration Number
Joseph A. Coppola	38,413	Jack L. Tribble	32,633

Direct all correspondence to: ☒ Customer Number or Bar Code Label 000210

Name	Joseph A. Coppola				
Address	Merck & Co., Inc. - Patent Department				
Address	P.O. Box 2000, RY60-30				
City	Rahway	**State**	NJ	**ZIP**	07065-0907
Country	USA	**Telephone**	(732)594-673	**Fax**	(732)594-4720

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. 1001 and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Name of Sole or First Inventor: ☐ A petition has been filed for this unsigned inventor

Given Name (first and middle [if any])	Family Name or Surname
KONSTANTIN	PETRUKHIN

Inventor's Signature					**Date**		
Residence: City	Collegeville	**State**	PA	**Country**	US	**Citizenship**	RU
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	**State**	NJ	**ZIP**	07065-0907		

☒ Additional inventors are being named on the 1 supplemental Additional Inventors(s) sheet(s) PTO/SB/02A attached hereto.

DECLARATION AND POWER OF ATTORNEY

ADDITIONAL INVENTOR(S)
Supplemental Sheet

Name of Additional Joint Inventor, if any:		☐ A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
C. THOMAS				CASKEY			
Inventor's Signature	[signature]					**Date**	6/19/01
Residence: City	Lansdale	**State**	PA	**Country**	US	**Citizenship**	US
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	**State**	NJ	**ZIP**	07065-0907		

Name of Additional Joint Inventor, if any:		☐ A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
Inventor's Signature						**Date**	
Residence: City		**State**		**Country**		**Citizenship**	
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	**State**	NJ	**ZIP**	07065-0907		

Name of Additional Joint Inventor, if any:		☐ A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
Inventor's Signature						**Date**	
Residence: City		**State**		**Country**		**Citizenship**	
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	**State**	NJ	**ZIP**	07065-0907		

Name of Additional Joint Inventor, if any:		☐ A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
Inventor's Signature						**Date**	
Residence: City		**State**		**Country**		**Citizenship**	
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	**State**	NJ	**ZIP**	07065-0907		

Please type a plus sign (+) inside this box → [+]

Approved for use through 9/30/2000. OMB 651-0032

SUBSTITUTE for PTO/SB/01 (12-97), DECLARATION FOR UTILITY OR DESIGN PATENT APPLICATION

DECLARATION AND POWER OF ATTORNEY FOR UTILITY OR DESIGN PATENT APPLICATION (37 CFR 1.63)

Attorney Docket Number	20267P
First Named Inventor	Petrukhin, et al.
COMPLETE IF KNOWN	
Application Number	09/806,088
Filing Date	March 26, 2001
Group Art Unit	
Examiner Name	

☐ Declaration Submitted with Initial Filing OR ☒ Declaration Submitted after Initial Filing (surcharge (37 CFR 1.16 (e)) required)

As a below named inventor, I hereby declare that:

My residence, post office address, and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

DELTA 6 FATTY DESATURASE

(Title of the Invention)

the specification of which

☐ is attached hereto

OR

☒ was filed on (MM/DD/YYYY) 03/26/2001 as United States Application Number or PCT International Application Number 09/806,088 and was amended on (MM/DD/YYYY) _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment specifically referred to above.

I acknowledge the duty to disclose to the Patent and Trademark Office all information known to me to be material to patentability as defined in 37 CFR 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. 119(a)-(d) or 365(b) of any foreign application(s) for patent or inventor's certificate, or 365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or of any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (MM/DD/YYYY)	Attorney Docket Number	Priority Claimed? YES	NO
PCT/US99/23253	PCT	10/09/1999	20267-PCT	☒	☐
				☐	☐
				☐	☐
				☐	☐

☐ Additional foreign application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto:

I hereby claim the benefit under 35 U.S.C. 119(e) of any United States provisional application(s) listed below.

Application Number(s)	Filing Date (MM/DD/YYYY)	Attorney Docket Number
60/103,760	10/09/1998	20267PV

Please type a plus sign (+) inside this box → [+]

Approved for use through 9/30/2000. OMB 651-0032

SUBSTITUTE for PTO/SB/01 (12-97), DECLARATION FOR UTILITY OR DESIGN PATENT APPLICATION

DECLARATION AND POWER OF ATTORNEY for Utility or Design Patent Application

I hereby claim the benefit under 35 U.S.C 120 of any United States application(s), or 365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of 35 U.S.C. 112, I acknowledge the duty to disclose information known to me to be material to patentability as defined in 37 CFR 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application or PCT Parent Application Number	Parent Filing Date (MM/DD/YYYY)	Parent Patent Number *(if applicable)*
60/103,760	10/09/1998	
pct/us99/23253	10/05/1999	

☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto.

As a named inventor, I hereby appoint, respectively and individually, as my attorneys or agents with full power of substitution and revocation, the following registered practitioner(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

☐ Customer Number [ ] → *Place Customer Number Bar Code Label here*

OR

[X] Registered practitioner(s) name/registration number listed below

Name	Registration Number	Name	Registration Number
Joseph A. Coppola	38,413	Jack L. Tribble	32,633

Direct all correspondence to: [X] Customer Number or Bar Code Label 000210

Name	Joseph A. Coppola						
Address	Merck & Co., Inc. - Patent Department						
Address	P.O. Box 2000, RY60-30						
City	Rahway			**State**	NJ	**ZIP**	07065-0907
Country	USA	**Telephone**	(732)594-673			**Fax**	(732)594-4720

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. 1001 and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Name of Sole or First Inventor: ☐ A petition has been filed for this unsigned inventor

Given Name (first and middle [if any])	Family Name or Surname
KONSTANTIN	PETRUKHIN

Inventor's Signature	Konstantin Petrukhin					**Date**	6/01/01
Residence: City	Collegeville	**State**	PA PA	**Country**	US	**Citizenship**	RU
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway			**State**	NJ	**ZIP**	07065-0907

[X] Additional inventors are being named on the 1 supplemental Additional Inventors(s) sheet(s) PTO/SB/02A attached hereto.

Please type a plus sign (+) inside this box → [+]

Approved for use through 9/30/2000. OMB 651-0032
SUBSTITUTE for PTO/SB/02A (3-97), Declaration (Additional Inventors)

DECLARATION AND POWER OF ATTORNEY

ADDITIONAL INVENTOR(S)
Supplemental Sheet

OIPE JC73 JUL 1 3 2001 PATENT & TRADEMARK OFFICE

Name of Additional Joint Inventor, if any:			☐ A petition has been filed for this unsigned inventor				
Given Name (first and middle [if any])			Family Name or Surname				
C. THOMAS			CASKEY				
Inventor's Signature					**Date**		
Residence: City	Lansdale	**State**	PA PA	**Country**	US	**Citizenship**	US
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	**State**	NJ	**ZIP**	07065-0907		

Name of Additional Joint Inventor, if any:			☐ A petition has been filed for this unsigned inventor				
Given Name (first and middle [if any])			Family Name or Surname				
Inventor's Signature					**Date**		
Residence: City		**State**		**Country**		**Citizenship**	
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	**State**	NJ	**ZIP**	07065-0907		

Name of Additional Joint Inventor, if any:			☐ A petition has been filed for this unsigned inventor				
Given Name (first and middle [if any])			Family Name or Surname				
Inventor's Signature					**Date**		
Residence: City		**State**		**Country**		**Citizenship**	
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	**State**	NJ	**ZIP**	07065-0907		

Name of Additional Joint Inventor, if any:			☐ A petition has been filed for this unsigned inventor				
Given Name (first and middle [if any])			Family Name or Surname				
Inventor's Signature					**Date**		
Residence: City		**State**		**Country**		**Citizenship**	
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	**State**	NJ	**ZIP**	07065-0907		